

FEATURE:
NAME/KEY: unsure
LOCATION: (653)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (656)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (660)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (661)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (662)
OTHER INFORMATION: Unidentified at time of filing
US-09-351-215-4

Query Match 93.1%; Score 27; DB 3; Length 888;
Best Local Similarity 83.3%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
Db 248 SSGPSI 253

RESULT 13
US-08-658-857B-16
; Sequence 16, Application US/08658857B
; Patent No. 6040435
; GENERAL INFORMATION:
; APPLICANT: Hancock, Robert E. W.
; APPLICANT: Karunaratne, Nedra
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: May 31, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/460,464
FILING DATE: June 2, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07420/014001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-658-857B-16

Query Match 89.7%; Score 26; DB 3; Length 20;
Best Local Similarity 83.3%; Pred. No. 46;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SSGPSL 6
Db 12 SSGPAL 17
RESULT 14
US-08-763-226C-16
; Sequence 16, Application US/08763226C
; Patent No. 6057291
; GENERAL INFORMATION:
; APPLICANT: Hancock, Robert E. W.
; APPLICANT: Karunaratne, Nedra
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/763,226C
FILING DATE: 10-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/658,857
FILING DATE: 31-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07420/014001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-763-226C-16

Query Match 89.7%; Score 26; DB 3; Length 20;
Best Local Similarity 83.3%; Pred. No. 46;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
Db 12 SSGPAL 17

RESULT 15
US-09-307-200-16
; Sequence 16, Application US/09307200
; Patent No. 6297215
; GENERAL INFORMATION:
; APPLICANT: Hancock, Robert E. W.
; APPLICANT: Karunaratne, Nedra
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla

```
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/307,200
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/763,226
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07420/014001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5039
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-307-200-16

Query Match      89.7%; Score 26; DB 3; Length 20;
Best Local Similarity 83.3%; Pred. No. 46;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSGPSL 6
Db      12 SSGPAL 17
```

Search completed: March 10, 2004, 09:28:53
Job time : 3.32296 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 26.7237 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-10

Perfect score: 186

Sequence: 1 FNNFTVSFWLRVPRKVSASHGSLHWSYGLRPX 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

```

1: /cgn2_6/prodata/2/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubaa/CTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubaa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubaa/US09_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/2/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/2/pubaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 185 | 99.5 | 34 | 9 | US-09-848-834A-10 |
| 2 | 185 | 99.5 | 50 | 9 | US-09-848-834A-18 |
| 3 | 119 | 64.0 | 194 | 14 | US-10-295-074-45 |
| 4 | 118 | 63.4 | 158 | 14 | US-10-297-942-12 |
| 5 | 116 | 62.4 | 285 | 14 | US-10-295-074-11 |
| 6 | 116 | 62.4 | 287 | 14 | US-10-295-074-13 |
| 7 | 114 | 61.3 | 158 | 14 | US-10-297-942-10 |
| 8 | 112 | 60.2 | 21 | 9 | US-09-843-548-3 |
| 9 | 112 | 60.2 | 21 | 9 | US-09-848-834A-4 |
| 10 | 112 | 60.2 | 21 | 9 | US-09-785-215-6 |
| 11 | 112 | 60.2 | 21 | 10 | US-09-405-986-2 |
| 12 | 112 | 60.2 | 21 | 14 | US-10-204-382-6 |
| 13 | 112 | 60.2 | 21 | 14 | US-10-339-522-3 |
| 14 | 112 | 60.2 | 21 | 14 | US-10-223-711-8 |
| 15 | 112 | 60.2 | 21 | 14 | US-10-223-809A-6 |

```

16 112 60.2 21 14 US-10-261-208-5
17 112 60.2 21 14 US-10-295-074-5
18 112 60.2 21 15 US-10-372-111-8
19 112 60.2 32 15 US-10-411-544-28
20 112 60.2 37 9 US-09-848-834A-14
21 112 60.2 158 14 US-10-297-942-2
22 112 60.2 158 14 US-10-297-942-18
23 112 60.2 158 14 US-10-297-942-20
24 112 60.2 194 14 US-10-295-074-47
25 112 60.2 285 14 US-10-295-074-9
26 112 60.2 287 14 US-10-295-074-15
27 112 60.2 441 15 US-10-452-024-180
28 112 60.2 441 15 US-10-452-024-183
29 112 60.2 444 15 US-10-452-024-185
30 112 60.2 451 15 US-10-452-024-186
31 112 60.2 452 15 US-10-452-024-184
32 112 60.2 463 10 US-09-816-467-2
33 112 60.2 469 15 US-10-452-024-182
34 112 60.2 472 15 US-10-452-024-181
35 112 60.2 514 14 US-10-295-074-49
36 112 60.2 514 14 US-10-295-074-51
37 112 60.2 514 14 US-10-295-074-59
38 112 60.2 517 14 US-10-295-074-53
39 112 60.2 573 15 US-10-452-024-177
40 112 60.2 605 14 US-10-130-973A-11
41 112 60.2 685 14 US-10-130-973A-9
42 112 60.2 882 14 US-10-130-973A-3
43 112 60.2 907 14 US-10-130-973A-5
44 112 60.2 1052 14 US-10-130-973A-17
45 112 60.2 1112 14 US-10-130-973A-16

```

ALIGNMENTS

RESULT 1

```

US-09-848-834A-10
; Sequence 10, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apelon Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/846,834A
; PRIORITY FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIORITY FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent version 3.0
; SEQ ID NO 10
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 947-967 of t
; OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylisin) linked by a spacer to a
; OTHER INFORMATION: Ino acid sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated phenylalanine
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(21)
; OTHER INFORMATION: Amino acids 947-967 of the Tetanus Toxoid Precursor
; OTHER INFORMATION: (Tentoxylisin)
; NAME/KEY: PEPTIDE
; LOCATION: (22)..(25)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (26)..(34)
; OTHER INFORMATION: Amino acids 2-10 of the human GnRH hormone
; NAME/KEY: MOD RES
; LOCATION: (34)..(34)
; OTHER INFORMATION: Amidated glycine or glycineamide

```

```
US-09-848-834A-10
Query Match          99.5%; Score 185; DB 9; Length 34;
Best Local Similarity 100.0%; Pred. No. 7.5e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 33
    |||||
Db 1 FNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 33

RESULT 2
US-09-848-834A-18
; Sequence 18, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphon Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the Tetanus
; OTHER INFORMATION: anox toxin precursor (Tentoxylisin) protein linked by a spacer to
; OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD RES
; LOCATION: (50)..(50)
; OTHER INFORMATION: Amidated glycine or glycine amide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(37)
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxin precursor (Tent
; OTHER INFORMATION: oxylisin
; NAME/KEY: PEPTIDE
; LOCATION: (38)..(41)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (42)..(50)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-18

Query Match          99.5%; Score 185; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.1e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 33
    |||||
Db 17 FNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 49

RESULT 3
US-10-295-074-46
; Sequence 46, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
; APPLICANT: Pharmex A/S

; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P1013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: hTNF with inserted tetanus toxoid P2 and P30 epitopes
; NAME/KEY: MUTAGEN
; LOCATION: (110)..(124)
; OTHER INFORMATION: Tetanus toxoid P2 epitope (SEQ ID NO: 2)
; FEATURE:
; NAME/KEY: MUTAGEN
; LOCATION: (125)..(145)
; OTHER INFORMATION: Tetanus toxoid P30 epitope (SEQ ID NO: 3)
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (2)..(109)
; OTHER INFORMATION: hTNF amino acids 1-108
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (146)..(194)
; OTHER INFORMATION: hTNF amino acids 109-157
US-10-295-074-46

Query Match          64.0%; Score 119; DB 14; Length 194;
Best Local Similarity 81.5%; Pred. No. 3.6e-08;
Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 FNNFTVFWLRVPKVSASHLEGPSLHW 27
    |||||
Db 125 FNNFTVFWLRVPKVSASHLEAKPW 151

RESULT 4
US-10-297-942-12
; Sequence 12, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
; APPLICANT: Ferring BV
; TITLE OF INVENTION: Solubilised Protein Vaccines
; FILE REFERENCE: P68445USO
; CURRENT APPLICATION NUMBER: US/10/297,942
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: PCT/DK01/00431
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: DK PA 2000 00966
; PRIOR FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-297-942-12

Query Match          63.4%; Score 118; DB 14; Length 158;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVFWLRVPKVSASHLEG 22
    |||||
Db 133 FNNFTVFWLRVPKVSASHLEG 154

RESULT 5
US-10-295-074-11
; Sequence 11, Application US/10295074
```

```
; Publication No. US20030185845A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P1013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: 2 human IL5 monomers joined by P2 and P30 epitopes
US-10-295-074-11

Query Match      62.4%; Score 116; DB 14; Length 285;
Best Local Similarity 91.7%; Pred. No. 1.4e-07;
Matches 22; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSWLRVPKVSASHLEGPS 24
Db 150 FNNFTVSWLRVPKVSASHLEIPT 173

RESULT 6
US-10-295-074-13
; Sequence 13, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P1013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Two human IL5 monomers joined by diglycine linker and including b
; OTHER INFORMATION: terminally positioned P30 and P2 epitopes
US-10-295-074-13

Query Match      62.4%; Score 116; DB 14; Length 287;
Best Local Similarity 91.7%; Pred. No. 1.4e-07;
Matches 22; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSWLRVPKVSASHLEGPS 24
Db 24 FNNFTVSWLRVPKVSASHLEIPT 47

RESULT 7
US-10-297-942-10
; Sequence 10, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
; APPLICANT: Ferring BV
; TITLE OF INVENTION: Solubilised Protein Vaccines
; FILE REFERENCE: P68445U0
; CURRENT APPLICATION NUMBER: US/10/297,942
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: PCT/DK01/00431
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: DK PA 2000 00966
; PRIOR FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
```

```
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-297-942-10

Query Match      61.3%; Score 114; DB 14; Length 158;
Best Local Similarity 70.6%; Pred. No. 1.4e-07;
Matches 24; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

QY 1 FNNFTVSWLRVPKVSASHLEG-----PSLH 26
Db 41 FNNFTVSWLRVPKVSASHLEQLVLFKGGQCPSTH 74

RESULT 8
US-09-943-548-3
; Sequence 3, Application US/09943548
; Patent No. US20020042364A1
; GENERAL INFORMATION:
; APPLICANT: Rittershaus, Charles W.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVIT
; FILE REFERENCE: TCS-411-IP US-1; TCS-411-IP US-2
; CURRENT APPLICATION NUMBER: US/09/943,548
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: helper T cell epitope of tetanus toxin
US-09-943-548-3

Query Match      60.2%; Score 112; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSWLRVPKVSASHLE 21
Db 1 FNNFTVSWLRVPKVSASHLE 21

RESULT 9
US-09-848-834A-4
; Sequence 4, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Tetanus bacillus
; NAME/KEY: PEPTIDE
; LOCATION: (1)-(21)
; OTHER INFORMATION: Amino acid sequence 947-967 of Tetanus
; OTHER INFORMATION: Toxoid Precursor (Tentoxylisin)
```

US-09-848-834A-4

Query Match 60.2%; Score 112; DB 9; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
 DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 10

US-09-785-215-6
 ; Sequence 6, Application US/09785215
 ; Publication No. US20020187157A1

; GENERAL INFORMATION:
 ; APPLICANT: JENSEN, Martin Roland et al.
 ; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID
 ; FILE REFERENCE: 3631-0107P
 ; CURRENT APPLICATION NUMBER: US/09/785,215
 ; CURRENT FILING DATE: 2001-02-20
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 21
 ; TYPE: PRT
 ; ORGANISM: Clostridium tetani

US-09-785-215-6

Query Match 60.2%; Score 112; DB 9; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
 DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 11

US-09-405-986-2
 ; Sequence 2, Application US/09405986
 ; Publication No. US20030157115A1

; GENERAL INFORMATION:
 ; APPLICANT: BAY, Sylvie
 ; APPLICANT: CANTAUZENNE, Daniele
 ; APPLICANT: LECLERC, Claude
 ; APPLICANT: LO-MAN, Richard
 ; TITLE OF INVENTION: MULTIPLE ANTIGEN GLYCOPEPTIDE CARBOHYDRATE, VACCINE
 ; FILE REFERENCE: 1341 US 3565

; CURRENT APPLICATION NUMBER: US/09/405,986

; CURRENT FILING DATE: 1999-09-27

; EARLIER APPLICATION NUMBER: 60/041,726

; EARLIER FILING DATE: 1997-03-27

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 21

; TYPE: PRT

; ORGANISM: Clostridium tetani

US-09-405-986-2

Query Match 60.2%; Score 112; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
 DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 12

US-10-204-362-6

; Sequence 6, Application US/10204362
 ; Publication No. US20030086938A1

; GENERAL INFORMATION:

; APPLICANT: M&E Biotech A/S

; TITLE OF INVENTION: No. US20030086938A1 Method For Down-Regulation Of Amyloid

; FILE REFERENCE: 3631-0120P

; CURRENT APPLICATION NUMBER: US/10/204,362

; CURRENT FILING DATE: 2002-08-16

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn Ver. 3.0

; SEQ ID NO 6

; LENGTH: 21

; TYPE: PRT

; ORGANISM: Clostridium tetani

US-10-204-362-6

Query Match 60.2%; Score 112; DB 14; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
 DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 13

US-10-339-522-3

; Sequence 3, Application US/10339522
 ; Publication No. US20030108559A1

; GENERAL INFORMATION:

; APPLICANT: Rittershaus, Charles W.

; APPLICANT: Thomas, Lawrence J.

; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY

; FILE REFERENCE: TCS-411-IP US-3

; CURRENT APPLICATION NUMBER: US/10/339,522

; CURRENT FILING DATE: 2003-01-08

; PRIOR APPLICATION NUMBER: 08/432,483

; PRIOR FILING DATE: 1995-05-01

; PRIOR APPLICATION NUMBER: PCT/US96/06147

; PRIOR FILING DATE: 1996-05-01

; PRIOR APPLICATION NUMBER: 08/945,289

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 09/943,334

; PRIOR FILING DATE: 2001-08-30

; PRIOR APPLICATION NUMBER: 09/943,548

; PRIOR FILING DATE: 2001-08-30

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 21

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: helper T cell epitope of tetanus toxin

US-10-339-522-3

Query Match 60.2%; Score 112; DB 14; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
 DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 14

US-10-223-711-8

; Sequence 8, Application US/10223711
 ; Publication No. US20030113344A1

; GENERAL INFORMATION:

; APPLICANT: Bakaletz, Lauren O.

; APPLICANT: Kaumaya, Pravin T.P.

; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides

```
; FILE REFERENCE: 18525/04058
; CURRENT APPLICATION NUMBER: US/10/223,711
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 09/148,711
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 08/460,502
; PRIOR FILING DATE: 1995-06-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-10-223-711-8
```

```
Query Match      60.2%; Score 112; DB 14; Length 21;
Best Local Similarity 100.0%; Pred.No. 3e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FNNFTVSFWLRVPKVSASHLE 21
        |||||
DB      1 FNNFTVSFWLRVPKVSASHLE 21
```

```
RESULT 15
US-10-223-809A-6
; Sequence 6, Application US/10223809A
; Publication No. US20030157117A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; APPLICANT: Rasmussen, Peter Birk et al.
; TITLE OF INVENTION: No. US20030157117A1 Method for Down-Regulation of Amyloid
; FILE REFERENCE: 674542-2008
; CURRENT APPLICATION NUMBER: US/10/223,809A
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: US 60/337,543
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: US 60/373,027
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: DE 2001 01231
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: DE 2002 0058
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-10-223-809A-6
```

```
Query Match      60.2%; Score 112; DB 14; Length 21;
Best Local Similarity 100.0%; Pred.No. 3e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FNNFTVSFWLRVPKVSASHLE 21
        |||||
DB      1 FNNFTVSFWLRVPKVSASHLE 21
```

Search completed: March 10, 2004, 10:25:48
Job time : 26.7237 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 33.6031 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-10
Perfect score: 186
Sequence: 1 FNNFTVFWLRVPKVSASHLEPSLHWSYGLRPX 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 112 | 60.2 | 451 | 2 Q9LA13 | Q9LA13 clostridium |
| 2 | 112 | 60.2 | 1310 | 2 Q93N27 | Q93N27 clostridium |
| 3 | 62 | 33.3 | 1268 | 2 Q45851 | Q45851 clostridium |
| 4 | 61 | 32.8 | 1278 | 2 Q57236 | Q57236 clostridium |
| 5 | 58 | 31.2 | 361 | 2 Q45848 | Q45848 clostridium |
| 6 | 58 | 31.2 | 361 | 2 Q45846 | Q45846 clostridium |
| 7 | 58 | 31.2 | 441 | 2 Q9X708 | Q9X708 clostridium |
| 8 | 58 | 31.2 | 1291 | 2 Q9ZAJ8 | Q9ZAJ8 clostridium |
| 9 | 58 | 31.2 | 1291 | 2 Q93G71 | Q93G71 clostridium |
| 10 | 58 | 31.2 | 1291 | 2 Q933K0 | Q933K0 clostridium |
| 11 | 58 | 31.2 | 1291 | 2 Q08077 | Q08077 clostridium |
| 12 | 58 | 31.2 | 1291 | 2 Q8GR96 | Q8GR96 clostridium |
| 13 | 57 | 30.6 | 430 | 2 Q9XAV1 | Q9XAV1 pseudomonas |
| 14 | 57 | 30.6 | 502 | 16 Q9X8T9 | Q9X8T9 streptomyce |
| 15 | 56.5 | 30.4 | 367 | 2 Q45861 | Q45861 clostridium |
| 16 | 56.5 | 30.4 | 367 | 2 Q45862 | Q45862 clostridium |

| | | | | | |
|----|------|------|------|-----------|----------------------|
| 17 | 56.5 | 30.4 | 1252 | 2 Q8KZM3 | Q8KZM3 clostridium |
| 18 | 56.5 | 30.4 | 1255 | 2 Q9FAR6 | Q9FAR6 clostridium |
| 19 | 56 | 30.1 | 804 | 10 Q8L853 | Q8L853 arabidopsis |
| 20 | 56 | 30.1 | 812 | 10 Q64620 | Q64620 arabidopsis |
| 21 | 56 | 30.1 | 1251 | 2 Q9K395 | Q9K395 clostridium |
| 22 | 55 | 29.6 | 1280 | 2 Q9ZAJ5 | Q9ZAJ5 clostridium |
| 23 | 54.5 | 29.3 | 476 | 16 Q8PM68 | Q8PM68 xanthomonas |
| 24 | 54.5 | 29.3 | 1285 | 2 Q45967 | Q45967 clostridium |
| 25 | 54.5 | 29.3 | 1285 | 2 Q9LBR1 | Q9LBR1 clostridium |
| 26 | 54.5 | 29.3 | 1291 | 2 Q93HT3 | Q93HT3 clostridium |
| 27 | 54 | 29.0 | 91 | 13 Q9PRH0 | Q9PRH0 anguilla ja |
| 28 | 53 | 28.5 | 606 | 16 Q8XIK6 | Q8XIK6 clostridium |
| 29 | 52 | 28.0 | 51 | 16 Q88GE4 | Q88GE4 pseudomonas |
| 30 | 52 | 28.0 | 202 | 16 Q829M1 | Q829M1 streptomyce |
| 31 | 52 | 28.0 | 466 | 12 Q8TXJ9 | Q8TXJ9 tomato spot |
| 32 | 52 | 28.0 | 466 | 12 Q8TXK0 | Q8TXK0 tomato spot |
| 33 | 52 | 28.0 | 467 | 12 Q8TXK2 | Q8TXK2 tomato spot |
| 34 | 52 | 28.0 | 467 | 12 Q37369 | Q37369 tomato spot |
| 35 | 52 | 28.0 | 467 | 12 Q37367 | Q37367 tomato spot |
| 36 | 52 | 28.0 | 467 | 12 Q8JVL0 | Q8JVL0 tomato spot |
| 37 | 52 | 28.0 | 467 | 12 Q8TXK4 | Q8TXK4 tomato sapien |
| 38 | 51.5 | 27.7 | 317 | 4 Q9Y2R8 | Q9Y2R8 homo sapien |
| 39 | 51.5 | 27.7 | 345 | 16 Q8VRF4 | Q8VRF4 bradyrhizob |
| 40 | 51.5 | 27.7 | 381 | 2 Q8VKZ1 | Q8VKZ1 enterobacte |
| 41 | 51.5 | 27.7 | 503 | 8 Q7YKX8 | Q7YKX8 utricularia |
| 42 | 51.5 | 27.7 | 503 | 8 Q7YKX7 | Q7YKX7 utricularia |
| 43 | 51.5 | 27.7 | 520 | 4 Q8WVD0 | Q8WVD0 homo sapien |
| 44 | 51.5 | 27.7 | 1127 | 16 Q8G815 | Q8G815 bifidobacte |
| 45 | 51 | 27.4 | 130 | 2 Q93PY5 | Q93PY5 pseudomonas |

ALIGNMENTS

RESULT 1
Q9LA13
ID Q9LA13 PRELIMINARY; PRT; 451 AA.
AC Q9LA13; 2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-OCT-2000 (TREMREL. 25, Last annotation update)
DE Tetanus toxin (Fragment).
OS Clostridium tetani
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=20886;
RA He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;
RT "Fragment C of Tetanus Toxin."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF154828; AAF73267.1; -
DR HGSP; P04958; IABD.
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; IEA.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR02160; Kunitz_legume.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
FT NON TER
SQ SEQUENCE 451 AA; 51823 MW; 69A8C5F030B6CD8E CRC64;
Query Match 60.2%; Score 112; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVFWLRVPKVSASHLE 21
Db 83 FNNFTVFWLRVPKVSASHLE 103

RESULT 2
Q93N27


```

ID Q93N27 PRELIMINARY; PRT; 1310 AA.
AC Q93N27;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tetanus toxin (Fragment).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
ON NCBI_TaxID=1513;
RX SHUIN Z., Dianliang L.;
RA "Cloning and sequence analysis of tetanus toxin gene.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF389424; AAK72964.2; -.
DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008985; ConA_like lec_gl.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR002160; Kunitz legume.
DR InterPro; IPR000395; Peptidase M27.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOKILYSIN.
DR PRODOM; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00225; CRYSTALLIN BETAGAMMA; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON TER 1 1310
FT TER 1310 1310
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418E450 CRC64;

Query Match 60.2%; Score 112; DB 2; Length 1310;
Best Local Similarity 100.0%; Pred. No. 5e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPSVASHLE 21
Db 948 FNNFTVSFWLRVPSVASHLE 968

RESULT 3
Q45851 PRELIMINARY; PRT; 1268 AA.
AC Q45851;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neurotoxin type F.
GN BONT /F.
OS Clostridium baratii.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1561;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9325228; PubMed=8486245;
RA Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,
RA Richardson P.T.;
RT "Nucleotide sequence of the gene coding for Clostridium baratii type F
neurotoxin: Comparison with other clostridial neurotoxins.";
RL FEMS Microbiol. Lett. 108:175-182(1993).
DR EMBL; X68262; CA48329.1; -.
DR FIR; S33411; S33411.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.

```

```

DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008985; ConA_like lec_gl.
DR InterPro; IPR002160; Kunitz legume.
DR InterPro; IPR000395; Peptidase_M27.
DR InterPro; IPR006025; Pept_M27_BS.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOKILYSIN.
DR PRODOM; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 1268 AA; 145513 MW; 963040091AC15ED2 CRC64;

Query Match 33.3%; Score 62; DB 2; Length 1268;
Best Local Similarity 64.3%; Pred. No. 6.5;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPSVASHLE 14
Db 922 YQNFVSFWVRIPK 935

RESULT 4
Q57236 PRELIMINARY; PRT; 1278 AA.
AC Q57236; Q45863;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Botulinum neurotoxin type F (BONT/F protein).
GN BONT/F.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RA Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;
RA Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 635-1000 FROM N.A.
RX STRAIN=NCTC 1028; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP SEQUENCE OF 1-27 FROM N.A.
RC STRAIN=LANGELEND;
RA East A.K., Bhandari M., Hiem S., Collins M.D.;
RT "Analysis of the botulinum neurotoxin type F gene clusters in
proteolytic and nonproteolytic Clostridium botulinum and Clostridium
baratii.";
RL Curr. Microbiol. 37:262-268(1998).
DR EMBL; X81714; CAA57358.1; -.
DR EMBL; L35496; AAZ23210.1; -.
DR EMBL; X70821; CAA50152.1; -.
DR EMBL; X99064; CAA67512.1; -.
DR FIR; S48110; S48110.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.

```

```

DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR InterPro; IPR002160; Kunitz_legume.
DR InterPro; IPR003195; Peptidase_M27.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF01742; Peptidase_K27; 1.
DR PRINTS; PR00760; BONTOKILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin.
SQ SEQUENCE 1278 AA; 147073 MW; 14BE1318431D6918 CRC64;

Query Match 32.8%; Score 61; DB 2; Length 1278;
Best Local Similarity 57.1%; Pred. No. 9.2;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPK 14
Db 931 YQNFISFWVRPK 944

RESULT 5
Q45848 PRELIMINARY; PRT; 361 AA.
ID AC Q45848;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Botulinum neurotoxin type B (Fragment).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-type B.
RA MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL; X70819; CAA50150.1; -.
DR HSSP; P10845; 3BTA.
DR GO: GO:0015070; F:toxin activity; IEA.
DR InterPro; IPR008985; ConA_like_lec_gl.
FT NON_TER 361
FT NON_TER 361
SQ SEQUENCE 361 AA; 42131 MW; A2E0FFFC81F9533D CRC64;

Query Match 31.2%; Score 58; DB 2; Length 361;
Best Local Similarity 64.3%; Pred. No. 6;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPK 14
Db 290 FLDFSVFWIRPK 303

RESULT 6
Q45846 PRELIMINARY; PRT; 361 AA.
ID AC Q45846;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Botulinum neurotoxin type B (Fragment).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-type B.
RA MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL; X70814; CAA50145.1; -.
DR HSSP; P10845; 3BTA.
DR GO: GO:0015070; F:toxin activity; IEA.
DR InterPro; IPR008985; ConA_like_lec_gl.
KW Neurotoxin.
FT NON_TER 361
FT NON_TER 361
SQ SEQUENCE 361 AA; 42175 MW; 533EA98735CD98E1 CRC64;

Query Match 31.2%; Score 58; DB 2; Length 361;
Best Local Similarity 64.3%; Pred. No. 6;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPK 14
Db 290 FLDFSVFWIRPK 303

RESULT 7
Q9X708 PRELIMINARY; PRT; 441 AA.
ID AC Q9X708;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Botulinum neurotoxin type B (Fragment).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99343691; PubMed=10413679;
RA Lalli G., Herreros J., Osborne S.L., Montecucco C., Rossetto O.,
RA Schiavo G.;
RT "Functional characterisation of tetanus and botulinum neurotoxins
RT binding domains.";
RL J. Cell Sci. 112:2715-2724(1999).
DR EMBL; AJ242628; CAB43706.1; -.
DR HSSP; P10845; 3BTA.
DR GO: GO:0004866; F:endorpeptidase inhibitor activity; IEA.
DR GO: GO:0015070; F:toxin activity; IEA.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR InterPro; IPR002160; Kunitz_legume.
KW Neurotoxin.
FT NON_TER 441
FT NON_TER 441
SQ SEQUENCE 441 AA; 52772 MW; 721D0B468B8C95A4 CRC64;

Query Match 31.2%; Score 58; DB 2; Length 441;
Best Local Similarity 64.3%; Pred. No. 7.5;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPK 14
Db 81 FLDFSVFWIRPK 94

RESULT 8
Q9ZAJ8 PRELIMINARY; PRT; 1291 AA.
ID AC Q9ZAJ8;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

```

```
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Bort protein.
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 3281;
RX MEDLINE=98440323; PubMed=9767710;
RA Santos-Buelga J., Collins M.D., East A.K.;
RT "Characterization of the genes encoding the Botulinum neurotoxin
complex in a strain of clostridium botulinum producing type B & F
neurotoxins.";
RL Curr. Microbiol. 37:312-318(1998).
DR EMBL; Y13630; CAA73968.1; -.
DR HSP; P10845; 3BTA.
DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0009405; F:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR002160; Kunitz legume.
DR InterPro; IPR000395; Peptidase M27.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF01742; Peptidase M27; 1.
DR PRINTS; PR00760; BONTOKILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 1291 AA; 150840 MW; E4D3B0E46AB2E735 CRC64;

Query Match 31.2%; Score 58; DB 2; Length 1291;
Best Local Similarity 64.3%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSEWLRVVK 14
| : : : : :
Db 923 FLDFSVFWIRPK 936

RESULT 9
ID Q93G71 PRELIMINARY; PRT; 1291 AA.
AC Q93G71;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Neurotoxin type B
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1436;
RX "Characterization of six type A strains of Clostridium botulinum that
contain type B toxin gene sequences.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF295926; AAK97132.1; -.
DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0009405; F:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002160; Kunitz legume.
DR InterPro; IPR000395; Peptidase M27.
DR InterPro; IPR006025; Pept_M_Zn_BS.

Query Match 31.2%; Score 58; DB 2; Length 1291;
Best Local Similarity 64.3%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSEWLRVVK 14
| : : : : :
Db 923 FLDFSVFWIRPK 936

RESULT 9
ID Q93G71 PRELIMINARY; PRT; 1291 AA.
AC Q93G71;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Neurotoxin type B
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1436;
RX "Characterization of six type A strains of Clostridium botulinum that
contain type B toxin gene sequences.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF295926; AAK97132.1; -.
DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0009405; F:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002160; Kunitz legume.
DR InterPro; IPR000395; Peptidase M27.
DR InterPro; IPR006025; Pept_M_Zn_BS.
```

```
DR Pfam; PF01742; Peptidase M27; 1.
DR PRINTS; PR00760; BONTOKILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 1291 AA; 150824 MW; D7CA07BAE2EB8CD2 CRC64;

Query Match 31.2%; Score 58; DB 2; Length 1291;
Best Local Similarity 64.3%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSEWLRVVK 14
| : : : : :
Db 923 FLDFSVFWIRPK 936

RESULT 10
ID Q933K0 PRELIMINARY; PRT; 1291 AA.
AC Q933K0;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Type B cryptic neurotoxin.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=593; and 588;
RA Kirma N., Ferreira J.L., Baumstark B.R.;
RT "Characterization of six type A strains of Clostridium botulinum that
contain type B toxin gene sequences.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF300466; AAL11499.1; -.
DR EMBL; AF300465; AAL11498.1; -.
DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0009405; F:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR002160; Kunitz legume.
DR InterPro; IPR000395; Peptidase M27.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF01742; Peptidase M27; 1.
DR PRINTS; PR00760; BONTOKILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 1291 AA; 150843 MW; 7AC1737B0FA5A151 CRC64;

Query Match 31.2%; Score 58; DB 2; Length 1291;
Best Local Similarity 64.3%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSEWLRVVK 14
| : : : : :
Db 923 FLDFSVFWIRPK 936

RESULT 11
ID Q08077 PRELIMINARY; PRT; 1291 AA.
AC Q08077;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE BONT/B.
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
```

```

DR PRODOM; PD001963; Bontoxilysin; 1
DR PROSITE; PS00142; ZINC PROTEASE; 1
SQ SEQUENCE 1291 AA; 150374 MW; 0227CAEFAF58504D CRC64;

Query Match 31.2%; Score 58; DB 2; Length 1291;
Best Local Similarity 64.3%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENNFTVSFWLRYPK 14
      :|||:||||:|
Db 923 FLDFSVFWIRPK 936

RESULT 13
Q9XAV1 PRELIMINARY; PRT; 430 AA.
ID Q9XAV1 AC Q9XAV1
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alkane 1-monooxygenase (EC 1.14.15.1).
DE ALKB.
GN GN
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHA0;
RX MEDLINE=21128526; PubMed=11207749;
RA Smith T.H.M., Roethlisberger M., Witholt B., Van Beilen J.B.;
RT "Molecular screening for alkane hydroxylase genes in Gram-negative and
RL Gram-positive bacteria.";
RT Environ. Microbiol. 1:307-317 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CHA0;
RA Smith T.H.M.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ009579; CAB51045.2; -.
DR GO; GO:0018683; F:camphor 5-monooxygenase activity; IEA.
DR InterPro; IPR005804; FA_desat_fam.
DR Pfam; PF00487; FA_desaturase_1.
DR Monooxygenase; Oxdoreductase.
SQ SEQUENCE 430 AA; 48337 MW; 048B950980783E86 CRC64;

Query Match 30.6%; Score 57; DB 2; Length 430;
Best Local Similarity 35.9%; Pred. No. 10;
Matches 14; Conservative 4; Mismatches 7; Indels 14; Gaps 2;

QY 7 SFWLRYPKV-----SASHLEG-----PSLHWSYGL 31
      :|||:||||:|
Db 224 SFWSFLPTVWFSLSSAWHLESQRLEKLGILPTLHWKGV 262

RESULT 14
Q9X8T8 PRELIMINARY; PRT; 502 AA.
ID Q9X8T8 AC Q9X8T8
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative transmembrane protein.
DE SC03902 OR SH24.24.
GN GN
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3 (2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,

```

Search completed: March 10, 2004, 09:25:33
Job time : 35.6031 secs

| Result No. | Score | Query | | | DB | ID | Description |
|------------|-------|-------|--------|---|-----------|---------------------|-------------|
| | | Match | Length | % | | | |
| 1 | 185 | 99.5 | 34 | 5 | AAU11421 | AAu11421 Synthetic | |
| 2 | 185 | 99.5 | 50 | 5 | AAU11429 | AAu11429 Synthetic | |
| 3 | 152 | 81.7 | 32 | 2 | AAR62702 | AAR62702 LHRH-cont | |
| 4 | 121.5 | 65.3 | 750 | 3 | AAy92633 | AAy92633 Mutant hu | |
| 5 | 121 | 65.1 | 537 | 7 | ABR82481 | ABR82481 Truncated | |
| 6 | 121 | 65.1 | 708 | 7 | ABR82479 | ABR82479 Modified | |
| 7 | 121 | 65.1 | 717 | 7 | ABR82478 | ABR82478 Modified | |
| 8 | 119 | 64.0 | 194 | 6 | AAO30488 | AAO30488 Human TNF | |
| 9 | 118.5 | 63.7 | 31 | 3 | AAy92655 | AAy92655 PSMep012 | |
| 10 | 118 | 63.4 | 158 | 2 | AAW81336 | AAW81336 TNF30-5, | |
| 11 | 118 | 63.4 | 158 | 5 | ABBO7279 | ABBO7279 Human TNF | |
| 12 | 117 | 62.9 | 750 | 3 | AAy92627 | AAy92627 Mutant hu | |
| 13 | 117 | 62.9 | 750 | 3 | AAy92656 | AAy92656 Mutant hu | |
| 14 | 116 | 62.4 | 188 | 3 | AAy84423 | AAy84423 An oestrog | |
| 15 | 116 | 62.4 | 285 | 3 | AAO30458 | AAO30458 hIL5-P2-P | |
| 16 | 116 | 62.4 | 287 | 6 | AAO30459 | AAO30459 hIL5,36 v | |
| 17 | 116 | 62.4 | 1807 | 4 | AAAB85697 | AAAB85697 Recombina | |
| 18 | 116 | 62.4 | 2028 | 4 | AAAB5698 | AAAB5698 Recombina | |
| 19 | 115.5 | 62.1 | 693 | 3 | AAy92647 | AAy92647 Mutant hu | |
| 20 | 115.5 | 62.1 | 693 | 3 | AAy92648 | AAy92648 Mutant hu | |
| 21 | 115.5 | 62.1 | 750 | 3 | AAy92645 | AAy92645 Mutant hu | |
| 22 | 115.5 | 62.1 | 750 | 3 | AAy92646 | AAy92646 Mutant hu | |
| 23 | 115 | 61.8 | 750 | 3 | AAy92634 | AAy92634 Mutant hu | |
| 24 | 114.5 | 61.6 | 109 | 4 | AAB30149 | AAB30149 Growth di | |
| 25 | 114 | 61.3 | 158 | 2 | AAW81333 | AAW81333 TNF30-2, | |

04-MAY-2001: 2001WO-US014363.

PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 XX Ladd AE, Wang CY, Zamb T;
 XX WPI; 1994-357910/44.
 XX Immunogenic luteinising hormone releasing hormone peptide(s) - that
 PT suppress LHRH activity in males and females.
 XX Claim 8; Page 84; 213pp; English.
 XX Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The stimulator
 CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
 CC immune response to the coupled peptide in members of a heterogeneous
 CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
 CC sequence from the invasive protein of Yersinia. Spacer amino acid
 CC sequences (e.g. Gly-Gly) can be provided between the invasive and Th
 CC domains and between the immune stimulator and hapten components. When the
 CC hapten is LHRH, then optionally the invasive domain can be omitted from
 CC the immune stimulator component. The present sequence represents an LHRH-
 CC containing, invasive-free immunogenic peptide as above which can be used
 CC as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-
 CC dependent carcinoma, prostatic carcinoma, testicular carcinoma,
 CC endometriosis, benign uterine tumours, recurrent functional ovarian
 CC cysts, (severe) premenstrual syndrome or oestrogen-dependent breast
 CC cancer, or for induction of infertility. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX SQ Sequence 32 AA;
 SQ Query Match 81.7%; Score 152; DB 2; Length 32;
 Best Local Similarity 87.9%; Pred. No. 5.6e-15;
 Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
 QY 1 FNNFTVSWLRVPKVSASHLEGPSSLHWSYGLRP 33
 DB 3 FNNFTVSWLRVPKVSASHLE----HWSYGLRP 31
 RESULT 4
 AAY92633
 ID AAY92633 standard; protein; 750 AA.
 AC AAY92633;
 XX 10-AUG-2000 (first entry)
 DE Mutant human prostate specific membrane antigen construct, hPSM1.10.
 XX Prostate specific membrane antigen; immunogenized construct; mutant;
 KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
 KW prostate cancer; cell-associated peptide antigen; foreign epitope.
 XX Homo sapiens.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Peptide 24..38
 FT /label= P2
 FT /note= "foreign epitope"
 FT Peptide 673..693
 FT /label= P30
 FT /note= "foreign epitope"
 XX WO200020027-A2.
 PN 13-APR-2000.
 XX 05-OCT-1999; 99WO-DK000525.

XX 05-OCT-1998; 98DK-00001261.
 PR 20-OCT-1998; 98US-0105011P.
 XX (MEBI-) M & E BIOTECH AS.
 XX Steinaa L, Mouriteen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 XX WPI; 2000-349917/30.
 DR Inducing immune responses to weakly immunogenic, tumor associated peptide
 PT antigens for the treatment of breast and prostate cancer.
 XX Example 1; Page; 220pp; English.
 XX AAY92627-49 are mutant immunogenized human prostate specific membrane
 CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
 CC The immunogenic analogues of PSM can be used in the claimed method as an
 CC autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
 CC binding regions and cysteine residues involved in disulfide bonds are
 CC preserved in the immunogenized forms. The method is used for inducing
 CC immune responses against weakly immunogenic cell-associated peptide
 CC antigens (PA) such as those associated with cancers (self-proteins), e.g.
 CC human prostate specific membrane antigen (PSM), heregulin 2 (her2) and/or
 CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
 CC simultaneous presentation by antigen producing cells (APCs) of the
 CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
 CC group derived from the PA and/or at least 1 B-cell group derived from the
 CC cell-associated PA; and (2) at least 1 first T helper cell group which is
 CC foreign to the animal. Analogues of human PSM, human Her2 and
 CC human/murine FGF8b comprising a substantial part of all known and
 CC predicted CTL and B-cell epitopes of the respective PA and including at
 CC least one foreign T helper epitope are also claimed. The method is used
 CC to treat prostate, prostate/breast or breast cancer when the PA is human
 CC PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
 CC from the wild type human PSM (AAY92619), which appears on pages 184-187
 CC of the specification
 XX SQ Sequence 750 AA;
 SQ Query Match 65.3%; Score 121.5; DB 3; Length 750;
 Best Local Similarity 86.2%; Pred. No. 6.5e-09;
 Matches 25; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 1 FNNFTVSWLRVPKVSASHLEGPSSLHWSY 29
 DB 673 FNNFTVSWLRVPKVSASHLE-PSSHNY 700
 RESULT 5
 ABR82481
 ID ABR82481 standard; protein; 537 AA.
 AC ABR82481;
 XX 20-NOV-2003 (first entry)
 DE Truncated human CEA-TT P2 and P30 epitopes.
 XX CEA; immune response; carcinoembryonic antigen; antigen presenting cell;
 KW APC; cytostatic; vaccine; human; tetanus toxoid; p2; p30; antigen.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Peptide 1..34
 FT /note= "signal peptide"
 FT Protein 35..537
 FT /note= "mature protein"
 XX WO2003059379-A2.

DR N-PSDB; ACP35964.

XX Inducing an immune response in humans against autologous carcinoembryonic

PT antigen (CEA) comprises administering a modified CEA polypeptide, a

PT nucleic acid encoding the polypeptide, or a microorganism expressing the

PT polypeptide.

XX Disclosure; Page 114-117; 140pp; English.

XX The invention relates to inducing an immune response against autologous

CC carcinoembryonic antigen (CEA) in an animal, e.g. human. The method

CC involves effecting uptake and processing by antigen presenting cells

CC (APCs) in the animal of at least 1 modified CEA polypeptide or of a

CC nucleic acid encoding the modified CEA polypeptide or of a microorganism

CC or virus expressing the modified CEA polypeptide to induce a CTL response

CC and an antibody response that targets the autologous CEA. The method is

CC useful in immunizing actively against diseases characterized by cells

CC that express CEA. The present sequence represents a modified human CEA

CC polypeptide that has tetanus toxoid (TT) P2 and P30 epitopes introduced

CC in its sequence

XX Sequence 717 AA;

SQ

Query Match 65.1%; Score 121; DB 7; Length 717;

Best Local Similarity 95.8%; Pred. No. 7.4e-09;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLEGTS 24

Db 693 FNNFTVSFWLRVPKVSASHLEGTS 716

RESULT 8

AAO30488

ID AAO30488 standard; protein; 194 AA.

AC AAO30488;

XX 22-SEP-2003 (first entry)

DT Human TNFalpha variant, TNF34-P2-P30.

DE

XX Multimeric protein; interleukin 5; IL5; TNFalpha; inflammatory disease;

KW tumour necrosis factor alpha; gene therapy; arthritis; human; mutant;

KW mutant; variant; tetanus toxoid; epitope.

XX Homo sapiens.

OS Unidentified.

OS Chimeric.

XX Key Location/Qualifiers

FT Region 2..109

FT /note= "Human TNF"

FT Region 110..124

FT /note= "Tetanus toxoid P2 epitope"

FT Region 125..145

FT /note= "Tetanus toxoid P30 epitope"

FT Region 146..194

FT /note= "Human TNF"

XX WO2003042244-A2.

PN

XX 22-MAY-2003.

PD

XX 15-NOV-2002; 2002WO-DK000764.

PF

XX 16-NOV-2001; 2001DK-00001702.

PR

XX 16-NOV-2001; 2001US-0331575P.

PR

XX (PHAR-) PHARMEXA AS.

PA (KLYS/) KLYSNER S.

PA (NIEL/) NIELSEN F S.

PA (BRAT/) BRATT T.

PA (VOLD/) VOLDORGB B.

PA (MOUR/) MOURITSEN S.

XX Klynsner S, Nielsen FS, Bratt T, Voldborg B, Mouritsen S;

XX WPI; 2003-449558/42.

XX New immunogenic analogue of a polymeric protein, useful for preparing a

PT composition for treating inflammatory diseases e.g. arthritis.

XX Claim 23; Page 158; 196pp; English.

XX The invention relates to immunogenic analogues of multimeric proteins

CC such as immunogenic variants of interleukin 5 (IL5) and tumour necrosis

CC factor alpha (TNF, TNFalpha) and methods for production of immunogenic

CC analogues. The immunogenic analogue is useful for preparing a composition

CC for treating inflammatory diseases, e.g., arthritis. It is also used in

CC gene therapy. The present sequence is human TNFalpha variant protein with

CC an inserted tetanus toxoid P2 and P30 epitopes. This sequence is used to

CC illustrate the method of the invention

XX Sequence 194 AA;

SQ

Query Match 84.0%; Score 119; DB 6; Length 194;

Best Local Similarity 81.5%; Pred. No. 3.2e-09;

Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLEGPSLHW 27

Db 125 FNNFTVSFWLRVPKVSASHLEAKPW 151

RESULT 9

AAAY92655

ID AAY92655 standard; peptide; 31 AA.

XX AC AAY92655;

XX 10-AUG-2000 (first entry)

DT PSMpep012 - P30 inserted in hPSM insertion position 10.

DE

XX Foreign epitope; P2; prostate specific membrane antigen; vaccination;

KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;

KW prostate cancer; cell-associated peptide antigen.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 6..26

FT /label= P30

XX WO2000020027-A2.

PN

XX 13-APR-2000.

PD

XX 05-OCT-1999; 99WO-DK000525.

PF

XX 05-OCT-1998; 98DX-00001261.

PR

XX 20-OCT-1998; 98US-0105011P.

PR

XX (MEBI-) M & E BIOTECH AS.

PA Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;

PI Gautam A, Birk P, Karlsson G;

XX WPI; 2000-349917/30.

XX Inducing immune responses to weakly immunogenic, tumor associated peptide

PT antigens for the treatment of breast and prostate cancer.

XX Example 1; Page 118; 220pp; English.

XX AAY2650-55 are peptides designed which correspond to the P2 and P30
 CC epitopes with 5 flanking human prostate specific membrane antigen (hPSM)
 CC amino acids in each end. The flanking amino acids correspond to the
 CC epitope insertion sites 6, 8 and 10. The peptides will be used in, e.g. T
 CC cell proliferation assays, but also for ELISA or other in vitro assays.
 CC The claims detail a method for inducing immune responses against weakly
 CC immunogenic cell-associated peptide antigens (PA) such as those
 CC associated with cancers (i.e. self-proteins) for example, hPSM,
 CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The method
 CC comprises effecting simultaneous presentation by antigen producing cells
 CC (APCs) of the animals immune system of: (1) at least 1 CTL (cytotoxic T-
 CC lymphocyte) group derived from the PA and/or at least 1 B-cell group
 CC derived from the cell-associated PA; and (2) at least 1 first T helper
 CC cell group which is foreign to the animal. Analogues of human PSM, human
 CC Her2 and human/murine FGF8b comprising a substantial part of all known
 CC and predicted CTL and B-cell epitopes of the respective PA and including
 CC at least one foreign T helper epitope (e.g. P2 and/or P30) are also
 CC claimed. The method is used to treat prostate, prostate/breast or breast
 CC cancer when the PA is human PSM, FGF8b and Her2, respectively
 XX SQ

Sequence 31 AA;

Query Match 63.7%; Score 118.5; DB 3; Length 31;
 Best Local Similarity 92.3%; Pred. No. 4.6e-10;
 Matches 24; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 FNNFTVGFWLKVPKVSASHLEGPSLH 26
 DB 6 FNNFTVGFWLKVPKVSASHLE-PSH 30

RESULT 10
 AAWS1336
 ID AAW81336 standard; protein; 158 AA.

AC AAW81336;
 XX
 DT 21-APR-1999 (first entry)
 XX
 DE TNF30-5, a TNF-alpha analogue.
 XX Human tumour necrosis factor-alpha; TNF-alpha; TNF-alpha analogue;
 KW vaccine; rheumatoid arthritis; Crohn's disease; ulcerative colitis;
 KW cancer; disseminated sclerosis; diabetes; psoriasis; osteoporosis;
 KW asthma.
 XX Synthetic.
 OS Homo sapiens.
 OS
 PN WO9846642-A1.
 XX
 PD 22-OCT-1998.
 XX
 PF 15-APR-1998; 98WO-DK000157.
 XX
 PR 15-APR-1997; 97DK-00000418.
 PR 24-APR-1997; 97US-0044187P.
 XX
 PA (FERR) FARM LAB FERRING AS.
 XX
 PI Jensen MR, Mouritsen S, Elsnar H, Dalum I;
 XX
 DR WPI; 1998-594561/50.
 DR N-PSDB; AAV68425.
 XX
 XX Modified human tumour necrosis factor-alpha - comprises immunodominant T
 PT cell epitope, useful in vaccines to treat or prevent TNF-associated
 PT diseases, e.g. cancer.
 XX
 PS Claim 15; Page 81-82; 134pp; English.
 XX
 CC The present sequence represents a modified human tumour necrosis factor-

CC alpha (TNF-alpha) analogue. The analogues have no residual TNF activity
 CC and are immunogenic in a large proportion of the human population (by
 CC using promiscuous epitopes). The TNF-alpha analogue is able to generate,
 CC in humans, neutralizing antibodies to wild-type human TNF alpha, has at
 CC least one fragment of TNF substituted by a peptide containing an
 CC immunodominant T-cell epitope, and at least one TNF-alpha B-cell epitope.
 CC The substitution causes a significant change in the amino acid sequence
 CC of any one of the strands in the front beta-sheet, any of the connecting
 CC loops or any of the B', I or D strands in the back beta-sheet. The TNF-
 CC alpha analogues are used as vaccines for treatment or prevention of
 CC diseases associated with excessive release or activity of TNF-alpha, e.g.
 CC rheumatoid arthritis, Crohn's disease, ulcerative colitis, cancer of any
 CC sort, disseminated sclerosis, diabetes, psoriasis, osteoporosis and
 CC asthma

Sequence 158 AA;

Query Match 63.4%; Score 118; DB 2; Length 158;
 Best Local Similarity 100.0%; Pred. No. 3.6e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVGFWLKVPKVSASHLEG 22
 DB 133 FNNFTVGFWLKVPKVSASHLEG 154

RESULT 11
 ABB07279
 ID ABB07279 standard; protein; 158 AA.

XX ABB07279;
 AC ABB07279;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human TNF-alpha analogue TNF30-5.

XX TNF-alpha; pharmaceutical; vaccine; self-protein; tumour necrosis factor;
 KW cetylpyridinium chloride; immunisation; antiinflammatory; antirheumatic;
 KW antiarthritic; antiulcer; cytostatic; antidiabetic; antipsoriatic;
 KW antiasthmatic; immunomodulator; neuroprotective; osteopathic; human;
 KW TNF30-5.

XX Homo sapiens.
 OS
 PN WO200197837-A1.
 XX
 PD 27-DEC-2001.

XX 20-JUN-2001; 2001WO-DK000431.
 XX
 PR 21-JUN-2000; 2000DK-00000966.
 XX
 PA (FERR) FERRING BV.

XX Olesen OP, Balchen T, Bouman MHEM;
 PI
 XX WPI; 2002-114542/15.
 DR N-PSDB; ABA94389.

XX Novel vaccine composition for prevention/treatment of self-protein-
 PT mediated pathology such as cancer, diabetes and asthma, comprises
 PT modified immunogenic self-protein and surfactant capable of acting as
 PT solubilizer.

Claim 21; Page 42-43; 55pp; English.

XX The invention provides a pharmaceutical vaccine composition (I) for the
 CC prevention or treatment of a self-protein-mediated pathology. The
 CC composition comprises at least one modified immunogenic self-protein
 CC (selected from modified TNF-alpha proteins) and a surfactant capable of
 CC acting as a solubilizer. (I) is useful for preventing or treating a self-
 CC protein-mediated pathology such as an inflammatory disease, rheumatoid
 CC arthritis, an inflammatory bowel disease (ulcerative colitis or Crohn's

CC disease), cancer, cachexia, multiple sclerosis, diabetes, psoriasis,
 CC osteoporosis or asthma. (I) is useful for inducing autoantibodies to a
 CC self-protein such as TNF (tumour necrosis factor)-alpha in a human
 CC subject. (II) comprising cetylpyridinium chloride as a component is useful
 CC for immunisation of a human subject or for treatment of a human
 CC inflammatory disease. The present sequence represents a human TNF-alpha
 CC analogue TNF30-5
 XX
 SQ Sequence 158 AA;

Query Match 63.4%; Score 118; DB 5; Length 158;
 Best Local Similarity 100.0%; Pred. No. 3.6e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNNFTVSWLRVPKVSASHLEG 22
 |||||
 DB 133 FNNFTVSWLRVPKVSASHLEG 154

RESULT 12
 AAY92627
 ID AAY92627 standard; protein; 750 AA.

XX AC AAY92627;

XX DT 10-AUG-2000 (first entry)

XX DE Mutant human prostate specific membrane antigen construct, hPSM1.1.

XX KW Prostate specific membrane antigen; immunogenized construct; mutant;
 KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
 KW prostate cancer; cell-associated peptide antigen; foreign epitope.

XX OS Homo sapiens.

XX OS Synthetic.

| | |
|---------------|--------------------------|
| XX FH Key | Location/Qualifiers |
| XX FT Peptide | 17..31 |
| XX FT | /label= P2 |
| XX FT | /note= "foreign epitope" |
| XX FT Peptide | 32..52 |
| XX FT | /label= P30 |
| XX FT | /note= "foreign epitope" |

XX PN WO200020027-A2.

XX PD 13-APR-2000.

XX PF 05-OCT-1999; 99WO-DK000525.

XX PR 05-OCT-1998; 98DK-00001261.

XX PR 20-OCT-1998; 98US-0105011P.

XX PA (MEBI-) M & E BIOTECH AS.

XX PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;

XX PI Gautam A, Birk P, Karlsson G;

XX DR WPI; 2000-349917/30.

XX PT Inducing immune responses to weakly immunogenic, tumor associated peptide
 PT antigens for the treatment of breast and prostate cancer.

XX PS Example 1; Page; 220pp; English.

XX CC AAY92627-49 are mutant immunogenized human prostate specific membrane
 CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
 CC The immunogenic analogues of PSM can be used in the claimed method as an
 CC autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
 CC binding regions and cysteine residues involved in disulfide bonds are
 CC preserved in the immunogenized forms. The method is used for inducing
 CC immune responses against weakly immunogenic cell-associated peptide
 CC antigens (PA) such as those associated with cancers (self-proteins), e.g.

CC human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
 CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
 CC simultaneous presentation by antigen producing cells (APCs) of the
 CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
 CC group derived from the PA and/or at least 1 B-cell group derived from the
 CC cell-associated PA; and (2) at least 1 first T helper cell group which is
 CC foreign to the animal. Analogues of human PSM, human Her2 and
 CC human/murine FGF8b comprising a substantial part of all known and
 CC predicted CTL and B-cell epitopes of the respective PA and including at
 CC least one foreign T helper epitope are also claimed. The method is used
 CC to treat prostate, prostate/breast or breast cancer when the PA is human
 CC PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
 CC from the wild type human PSM (AAY92619), which appears on pages 184-187
 CC of the specification

XX SQ Sequence 750 AA;

Query Match 62.9%; Score 117; DB 3; Length 750;

Best Local Similarity 95.7%; Pred. No. 3e-08;

Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FNNFTVSWLRVPKVSASHLEGP 23

DB 32 FNNFTVSWLRVPKVSASHLETP 54

RESULT 13

AAY92636

ID AAY92636 standard; protein; 750 AA.

XX AC AAY92636;

XX DT 10-AUG-2000 (first entry)

XX DE Mutant human prostate specific membrane antigen construct, hPSM1.5.

XX KW Prostate specific membrane antigen; immunogenized construct; mutant;
 KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
 KW prostate cancer; cell-associated peptide antigen; foreign epitope.

XX OS Homo sapiens.

XX OS Synthetic.

| | |
|---------------|--------------------------|
| XX FH Key | Location/Qualifiers |
| XX FT Peptide | 24..38 |
| XX FT | /label= P2 |
| XX FT | /note= "foreign epitope" |
| XX FT Peptide | 301..321 |
| XX FT | /label= P30 |
| XX FT | /note= "foreign epitope" |

XX PN WO200020027-A2.

XX PD 13-APR-2000.

XX PF 05-OCT-1999; 99WO-DK000525.

XX PR 05-OCT-1998; 98DK-00001261.

XX PR 20-OCT-1998; 98US-0105011P.

XX PA (MEBI-) M & E BIOTECH AS.

XX PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;

XX PI Gautam A, Birk P, Karlsson G;

XX DR WPI; 2000-349917/30.

XX PT Inducing immune responses to weakly immunogenic, tumor associated peptide
 PT antigens for the treatment of breast and prostate cancer.

XX PS Example 1; Page; 220pp; English.

XX CC AAY92627-49 are mutant immunogenized human prostate specific membrane

| | | |
|----|---------------------------|--|
| OS | Synthetic. | |
| OS | Clostridium tetani. | |
| OS | Mus musculus. | |
| XX | | |
| XX | Key | Location/Qualifiers |
| FT | Peptide | 1. .14 |
| FT | Protein | /note= "His tag" |
| FT | Protein | 15. .112 |
| FT | Protein | /note= "residues 158-255 of murine OPGL" |
| FT | Peptide | 113. .133 |
| FT | Protein | /note= "tetanus toxoid p30 epitope" |
| FT | Protein | 134. .188 |
| FT | Protein | /note= "residues 262-316 of murine OPGL" |
| XX | | |
| PN | WO200015807-A1. | |
| XX | | |
| PD | 23-MAR-2000. | |
| XX | | |
| XX | 13-SEP-1999; | 99WO-DK000481. |
| XX | | |
| PR | 15-SEP-1998; | 98DK-00001164. |
| PR | 02-OCT-1998; | 98US-0102896P. |
| XX | | |
| PA | (MEBI-) M & E BIOTECH AS. | |
| XX | | |

| | |
|-----------|---|
| RESULT 15 | |
| AAO30458 | |
| ID | AAO30458 standard; protein; 285 AA. |
| XX | |
| XX | AAO30458; |
| XX | |
| XX | 22-SEP-2003 (first entry) |
| XX | |
| DE | HLI5-P2-P30-HIL5 (hIL5.35) fusion construct protein. |
| XX | |
| XX | Multimeric protein; interleukin 5; IL5; TNFalpha; inflammatory disease; |
| KW | tumour necrosis factor alpha; gene therapy; arthritis; interleukin 5; |
| KW | IL5; epitope; human; tetanus toxoid; chimeric. |
| XX | |
| OS | Homo sapiens. |
| OS | Unidentified. |
| OS | Chimeric. |
| XX | |
| XX | Key |
| FT | Peptide |
| FT | 1. .19 |
| FT | /note= "Human IL5 leader peptide" |
| FT | Protein |
| FT | 20. .285 |
| XX | /note= "Mature hIL5.35 protein" |
| XX | |
| PN | WO2003042244-A2. |
| XX | |
| PD | 22-MAY-2003. |
| XX | |
| PF | 15-NOV-2002; 2002WO-DK000764. |
| XX | |
| XX | 16-NOV-2001; 2001DK-00001702. |
| PR | |
| PR | 16-NOV-2001; 2001US-0331575P. |
| XX | |
| XX | (PHAR-) PHARMEXA AS. |
| PA | (KLYS/) KLYSNER S. |
| PA | (NIEL/) NIELSEN P S. |
| PA | |

PA (BRAT/) BRATT T.
 PA (VOLD/) VOLDORG B.
 PA (MOUR/) MOURITSEN S.
 XX
 XX
 PI Klysner S, Nielsen FS, Bratt T, Voldborg B, Mouritsen S;
 XX
 XX WPI; 2003-449558/42.
 DR N-PSDB; AAL61294.
 DR
 XX
 XX
 PT New immunogenic analogue of a polymeric protein, useful for preparing a
 PT composition for treating inflammatory diseases e.g. arthritis.
 XX
 XX
 PS Claim 20; Page 112-113; 196pp; English.
 XX
 XX The invention relates to immunogenic analogues of multimeric proteins
 CC such as immunogenic variants of interleukin 5 (IL5) and tumour necrosis
 CC factor alpha (TNF, TNFalpha) and methods for production of immunogenic
 CC analogues. The immunogenic analogue is useful for preparing a composition
 CC for treating inflammatory diseases, e.g., arthritis. It is also used in
 CC gene therapy. The present sequence is a fusion construct which comprises
 CC 2 human interleukin 5 (IL5) monomers joined by tetanus toxoid epitopes
 CC P30 and P2. This sequence is used to illustrate the method of the
 CC invention
 XX
 SQ Sequence 285 AA;

Query Match 62.4%; Score 116; DB 6; Length 285;
 Best Local Similarity 91.7%; Pred. No. 1.4e-08;
 Matches 22; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSEFWLRVPKVSASHLEGPS 24
 |||||
 Db 150 FNNFTVSEFWLRVPKVSASHLEIPT 173

Search completed: March 10, 2004, 09:12:10
 Job time : 52.1984 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: March 10, 2004, 08:58:54 ; Search time 10.5837 seconds
(without alignments)
309.015 Million cell updates/sec

Title: US-09-848-834A-10
Perfect score: 186
Sequence: 1 FNFTVGFWRVPKVSASHLEGPGLHWSYGLRPX 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 112 | 60.2 | 1315 | 1 BTCLTN | tentoxilysin (EC 3 |
| 2 | 62 | 33.3 | 1368 | 2 S33411 | botulinum neurotox |
| 3 | 61.5 | 33.1 | 92 | 1 RHRTG | gonadoliberin prec |
| 4 | 61 | 32.8 | 366 | 2 S48110 | neurotoxin type F |
| 5 | 61 | 32.8 | 369 | 2 S48109 | neurotoxin type F |
| 6 | 61 | 32.8 | 1274 | 2 I40813 | neurotoxin - Clost |
| 7 | 61 | 32.8 | 1297 | 2 S32791 | neurotoxin - Clost |
| 8 | 59 | 31.7 | 1296 | 1 BTCLAB | bontoxilysin (EC 3 |
| 9 | 58 | 31.2 | 1291 | 1 A48940 | bontoxilysin (EC 3 |
| 10 | 58 | 31.2 | 1291 | 2 I40631 | non-proteolytic bo |
| 11 | 57.5 | 30.9 | 90 | 1 RHMSG | gonadoliberin prec |
| 12 | 57 | 30.6 | 502 | 2 T36589 | probable transmemb |
| 13 | 56.5 | 30.4 | 367 | 2 S48106 | neurotoxin type B |
| 14 | 56.5 | 30.4 | 1251 | 2 JH0256 | botulinum neurotox |
| 15 | 56.5 | 30.4 | 1252 | 2 S21178 | botulinum neurotox |
| 16 | 56 | 30.1 | 812 | 2 T01618 | hypothetical prote |
| 17 | 56 | 30.1 | 1296 | 2 I40645 | botulinum neurotox |
| 18 | 54.5 | 29.3 | 92 | 1 RHUG | gonadoliberin prec |
| 19 | 54.5 | 29.3 | 1285 | 2 S70582 | botulinum neurotox |
| 20 | 54.5 | 29.3 | 1291 | 2 S46431 | botulinum neurotox |
| 21 | 54.5 | 29.3 | 1291 | 2 A49777 | botulinum neurotox |
| 22 | 54 | 29.0 | 67 | 2 I78541 | gonadoliberin prec |
| 23 | 53.5 | 28.8 | 469 | 2 B37837 | probable alpha-am |
| 24 | 53.5 | 28.8 | 3122 | 2 T17202 | DNA-directed DNA p |
| 25 | 53 | 28.5 | 519 | 2 S78196 | probable maturase |
| 26 | 52 | 28.0 | 10 | 1 RHQGG | gonadoliberin - pi |
| 27 | 52 | 28.0 | 10 | 1 RHSHG | gonadoliberin - sh |
| 28 | 52 | 28.0 | 89 | 2 I51423 | gonadoliberin prec |
| 29 | 52 | 28.0 | 449 | 2 S23158 | nucleocapsid prote |

| | | | | | |
|----|------|------|------|----------|--------------------|
| 30 | 52 | 28.0 | 464 | 1 MNVUWC | nonstructural prot |
| 31 | 52 | 28.0 | 467 | 1 MNVUW1 | nonstructural prot |
| 32 | 51.5 | 27.7 | 91 | 2 J07333 | madaka-type gonado |
| 33 | 51.5 | 27.7 | 537 | 2 S78195 | cytochrome-c oxida |
| 34 | 51.5 | 27.7 | 573 | 2 S78197 | probable maturase |
| 35 | 51 | 27.4 | 292 | 2 S22441 | protein kinase (EC |
| 36 | 51 | 27.4 | 496 | 2 T38197 | probable myb-like |
| 37 | 51 | 27.4 | 659 | 2 P70453 | aconitase - Aquife |
| 38 | 51 | 27.4 | 944 | 2 T18627 | hypothetical prote |
| 39 | 51 | 27.4 | 1196 | 2 J01467 | toxin, nontoxic co |
| 40 | 51 | 27.4 | 1196 | 2 S46430 | botulinum neurotox |
| 41 | 50.5 | 27.2 | 438 | 2 T20882 | alpha galactosyltr |
| 42 | 50.5 | 27.2 | 1450 | 2 T30273 | hypothetical prote |
| 43 | 50 | 26.9 | 432 | 2 T04726 | hypothetical prote |
| 44 | 50 | 26.9 | 591 | 2 S04401 | spheroidene monoox |
| 45 | 50 | 26.9 | 836 | 2 J00248 | ATP-binding cass |

ALIGNMENTS

RESULT 1
BTCLTN
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N;Alternate names: tetanus neurotoxin
C;Species: Clostridium tetani
C;Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 03-Jun-2002
C;Accession: A25689; A25757; B25194; A60759; S69364
R;Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel,
EMBO J. 5, 2495-2502, 1986
A;Title: Tetanus toxin: Primary structure, expression in E. coli, and homology with bo
A;Reference number: A25689; MUID:87053814; PMID:3536478
A;Accession: A25689
A;Molecule type: DNA
A;Residues: 1-1315 <EIS>
A;Cross-references: GB:X04436; NID:940769; PIDN:CAA28033.1; PID:940770
R;Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A;Title: The complete nucleotide sequence of tetanus toxin.
A;Reference number: A25757; MUID:87040747; PMID:3774547
A;Accession: A25757
A;Molecule type: DNA
A;Residues: 1-1315 <FAI>
A;Cross-references: GB:X06214; NID:940773; PIDN:CAA29564.1; PID:940774
R;Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A;Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in
A;Reference number: A25194; MUID:86085672; PMID:3510187
A;Accession: A25194
A;Molecule type: DNA
A;Residues: 743-1315 <FA2>
A;Cross-references: GB:M12739; NID:9144920; PIDN:AAA23282.1; PID:9144921
A;Accession: B25194
A;Molecule type: protein
A;Residues: 865-894 <FA3>
R;Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A;Title: Isolation, purification, and characterization of fragment B, the NH-2-termina
A;Reference number: A60759; MUID:90035436; PMID:2478476
A;Accession: A60759
A;Molecule type: protein
R;Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A;Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
A;Reference number: J50098; MUID:89093918; PMID:2463305
R;Schliavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B
Nature 359, 832-835, 1992
A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteol
A;Reference number: S27125; MUID:93063293; PMID:1331807
A;Contents: annotation

R;de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
 Eur. J. Biochem. 229, 61-69, 1995
 A;Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
 A;Reference number: S69348; MUID:95262688; PMID:7744050
 A;Accession: S69348
 A;Molecule type: protein
 A;Residues: 2-31 <DB>
 C;Comment: The source of this protein was an extrachromosomal plasmid.
 C;Comment: The precursor is cleaved by endogenous proteinase activity to form light (fra
 dual chains are not toxic when separated). The amino end of the heavy chain (fragment B)
 C;Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to ganglionic
 C;Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized
 Presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of sy
 C;Function:
 A;Description: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in synap
 C;Superfamily: tetanus toxin
 C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
 F;2-457/Product: tentoxylisin light chain (fragment A) #status predicted <TTL>
 F;461-1315/Product: tentoxylisin heavy chain (fragment B.C) #status experimental <TTH>
 F;461-864/Domain: channel forming (fragment B) #status predicted <TXB>
 F;865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>
 F;233,237/Binding site: zinc (His) #status predicted
 F;234/Active site: Glu #status predicted

Query Match 60.2%; Score 112; DB 1; Length 1315;
 Best Local Similarity 100.0%; Pred. No. 2e-07;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVVKVSASHLE 21
 DB 947 FNNFTVSFWLRVVKVSASHLE 967

RESULT 2

botulinum neurotoxin type F - Clostridium barati

C;Species: Clostridium barati
 C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
 C;Accession: S33411; S31860
 R;Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.T.
 FEMS Microbiol. Lett. 108, 175-182, 1993
 A;Title: Nucleotide sequence of the gene coding for Clostridium barati type F neurotoxin
 A;Reference number: S33411; MUID:93252228; PMID:8486245
 A;Accession: S33411
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1268 <THO>
 A;Cross-references: EMBL:X68262; NID:949138; PIDN:CAA48329.1; PID:949139
 C;Superfamily: tetanus toxin
 C;Keywords: neurotoxin

Query Match 33.3%; Score 62; DB 2; Length 1268;
 Best Local Similarity 64.3%; Pred. No. 2.2;
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVVK 14
 DB 922 YQNFVSFWVRIPK 935

RESULT 3

RHRTG

Gonadoliberin precursor - rat

N;Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hormon
 N;Contains: Gonadoliberin; prolactin release-inhibiting factor
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
 C;Accession: A40147; B26173; A48410
 R;Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
 Mol. Endocrinol. 3, 1257-1262, 1989
 A;Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex
 A;Reference number: A40147; MUID:89384661; PMID:2476669
 A;Accession: A40147

A;Molecule type: DNA
 A;Residues: 1-92 <BON>
 A;Cross-references: GB:M31670; NID:G204447; PIDN:AAA1264.1; PID:G204448
 R;Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
 Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
 A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonado
 A;Reference number: A94090; MUID:86034338; PMID:2867548
 A;Accession: B26173
 A;Molecule type: mRNA
 A;Residues: 1-92 <AD>
 A;Cross-references: GB:M2579; NID:G204445; PIDN:AAA1263.1; PID:G204446
 R;Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Bialock, J.E.
 Cell. Mol. Neurobiol. 12, 447-454, 1992
 A;Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormon
 A;Reference number: A48410; MUID:93105480; PMID:1468115
 A;Accession: A48410
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-92 <MAI>
 A;Cross-references: GB:S50870; NID:G262059; PIDN:AAB24572.1; PID:G262060
 A;Experimental source: thymus
 A;Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBIP:121083)
 C;Genetics: 47/3; 79/3
 A;Introns: 47/3; 79/3
 C;Function:
 A;Description: stimulates pituitary secretion of lutropin and follitropin
 A;Note: Gonadoliberin-associated protein may have prolactin release inhibiting activity
 C;Superfamily: gonadoliberin
 C;Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid;
 F;1-23/Domain: signal sequence #status predicted <SIG>
 F;24-92/Product: progadoliberin #status predicted <PGN>
 F;24-33/Product: gonadoliberin #status predicted <GLN>
 F;37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
 F;24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicte
 F;33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gl

Query Match 33.1%; Score 61.5; DB 1; Length 92;
 Best Local Similarity 51.7%; Pred. No. 0.14;
 Matches 15; Conservative 3; Mismatches 4; Indels 7; Gaps 2;

QY 12 VPKVSAS-----HLEG-PSLHWSYGLRP 33
 DB 4 IPLKMAAVLLTVCLEGGCSQHSYGLRP 32

RESULT 4

S48110

neurotoxin type F - Clostridium botulinum (fragment)

C;Species: Clostridium botulinum
 C;Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
 C;Accession: S48110
 R;Campbell, K.D.; Collins, M.D.; East, A.K.
 J. Clin. Microbiol. 31, 2255-2262, 1993
 A;Title: Gene probes for identification of the botulin neurotoxin gene and specific i
 A;Reference number: S48103; MUID:94013372; PMID:8408542
 A;Accession: S48110
 A;Status: preliminary; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-366 <CAM>
 A;Cross-references: EMBL:X70821; NID:9407792; PIDN:CAA50152.1; PID:9407793
 C;Superfamily: tetanus toxin
 C;Keywords: neurotoxin

Query Match 32.8%; Score 61; DB 2; Length 366;
 Best Local Similarity 57.1%; Pred. No. 0.76;
 Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVVK 14
 DB 297 YQNFVSFWVRIPK 310

RESULT 5

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-369 'CAM'
A;Cross-references: EMBL:X70820; NID:G407790; PIDN:CAA50151.1; PID:G407791
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C;Superfamily: tetanus toxin

RESULT 8
BtCtAB
botoxilysin (EC 3.4.24.69) A precursor - Clostridium botulinum
N;Alternate names: botulinum neurotoxin type A
C;Species: Clostridium botulinum
C;Date: 31-Mar-1993 #sequence;revision 31-Mar-1993 #text;change 18-Jun-1999
C;Accession: A35294; S03492; S68220; A33401; A53984; A60025; A27000
R;Binz, T.; Kurazano, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990
A;Title: The complete sequence of botulinum neurotoxin type A and comparison with other
A;Reference number: A35294; MIMD:90264400; PMID:2160960

A:Reference number: A35294; MUID:90264400; PMID:2160960
A:Accession: A35294
A:Molecule type: DNA
A:Residues: 1-1296 <BIN>
A:Cross-references: GB:M30196; NID:9144864; PIDN:AAA23262.1; PID:9144865
A:Experimental source: strain 62A, subtype A
R:Thompson, D.E.; Brehm, J.K.; Cultran, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson, Eur. J. Biochem. 189, 73-81, 1990
A:Title: The complete amino acid sequence of the Clostridium botulinum type A neurotoxin
A:Reference number: S09432; MUID:90235864; PMID:2185020
A:Accession: S09432
A:Molecule type: DNA
A:Residues: 1,'Q',3-26,'V',28-1296 <THO>
A:Cross-references: EMBL:X50666; NID:540381; PIDN:CAA36289.1; PID:940382
A:Experimental source: NCTC 2916
R:Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K. FEBS Lett. 376, 41-44, 1995
A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin component
A:Reference number: S67988; MUID:95096783; PMID:8521962
A:Accession: S68220
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-12 <FUI>
A:Cross-references: EMBL:D67030; DBEJ:D50421; NID:92160224
R:Betley, M.J.; Somers, B.; DasGupta, B.R. Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989
A:Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-terminus
A:Reference number: A33401; MUID:89350959; PMID:2669749

A:Reference number: A33401; MUID:89350959; PMID:2669749
A:Accession: A33401
A:Molecule type: DNA
A:Residues: 1-35 <RET>
A:Cross-references: GB:M27892; NID:G144880; PIDN:AAR23289.1; PID:G551776
R:Gienez, J.A.; DasGupta, B.R.
J. Protein Chem. 12, 351-363, 1993
A>Title: Boculinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42.
A:Reference number: A53884; MUID:94000342; PMID:8397793

A.Accession: A53884
A.Status: preliminary
A.Molecule type: protein
A.Pesidues: 867-880:1148-1217,'Y',1219 <GIM>
A.Experimental source: Strain Hall
A.Note: sequence extracted from NCBI backbone (NCBIP:139159); sequence modified after R.DasGupta, B.R.; Dekleva, M.L.
Biochimie 72, 661-664, 1990
A.Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and ar
A.Reference number: A60025; MUID:91120847; PMID:2126206
A.Accession: A60025

A;Molecule type: protein
 A;Residues: 2-6/445-453,'X',455-457 <DAS1>
 R;DasGupta, B.R.; Foley, J.; Niece, R.
 Biochemistry 26, 4162, 1987
 A;Title: Partial sequence of the light chain of botulinum neurotoxin type A.
 A;Reference number: A27000
 A;Accession: A27000
 A;Molecule type: protein
 A;Residues: 2-47 <DAS2>
 R;Binz, T.; Blas, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.;
 J. Biol. Chem. 269, 1617-1620, 1994
 A;Title: Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.
 A;Reference number: A49708; MUID:94124495; PMID:8294407
 A;Contents: annotation
 C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synap
 C;Genetics:
 A;Gene: atx; bota
 C;Function:
 A;Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-associated
 C;Superfamily: tetanus toxin
 C;Keywords: disulfide bond; hydrolase; metalloproteinase; neurotoxin; transmembrane prot
 F;2-444/Product: bontoxilysin A light chain #status experimental <IGHT>
 F;445-1296/Product: bontoxilysin A heavy chain #status experimental <H>
 F;223,227/Binding site: zinc (His) #status predicted
 F;224/Active site: Glu #status predicted

Query Match 31.7%; Score 59; DB 1; Length 1296;
 Best Local Similarity 57.1%; Pred. No. 5.9; Mismatches 4; Indels 0; Gaps 0;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
 Db 938 YENFTSFWIRPK 951

RESULT 9
 A48940
 bontoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum
 N;Alternate names: botulinum neurotoxin type B (BoNT/B)
 C;Species: Clostridium botulinum
 C;Date: 19-Dec-1993 #sequence revision 18-Nov-1994 #text change 18-Jun-1999
 C;Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08573; S08574
 R;Wheelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Manton, N.P.
 Appl. Environ. Microbiol. 58, 2345-2354, 1992
 A;Title: Molecular cloning of the Clostridium botulinum structural gene encoding the typ
 A;Reference number: A48940; MUID:92384550; PMID:1514783
 A;Accession: A48940
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1291 <WHE>
 A;Cross-references: GB:M81186; NID:g144734; PID:AAA23211.1; PID:g144735
 A;Experimental source: type B, Danish
 A;Note: sequence extracted from NCBI backbone (NCBIN:112080, NCIP:112081); this publica
 R;Campbell, K.D.; Collins, M.D.; East, A.K.
 J. Clin. Microbiol. 31, 2255-2262, 1993
 A;Title: Gene probes for identification of the botulinum neurotoxin gene and specific id
 A;Reference number: S48103; MUID:94013372; PMID:8406542
 A;Accession: S48105
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 634-994 <CAM>
 A;Cross-references: EMBL:X70817; NID:g407782; PIDN:CAA50148.1; PID:g407783
 A;Experimental source: proteolytic type B, strain NCTC 7273
 R;Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
 submitted to the EMBL Data Library, April 1992
 A;Description: Partial amino acid sequence of botulinum neurotoxin type B and compariso
 A;Reference number: S21575
 A;Accession: S21575
 A;Molecule type: DNA
 A;Residues: 36-217,'G',219-224,'S',226-246 <SZA>
 A;Cross-references: EMBL:Z11934; NID:g40383; PIDN:CAA77991.1; PID:g40384
 R;Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars, R
 J. Biol. Chem. 267, 14721-14729, 1992

A;Title: Minimal essential domains specifying toxicity of the light chains of tetanus t
 A;Reference number: A42871; MUID:92340509; PMID:1634516
 A;Accession: A42871
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-313,'S',315-451 <KUR>
 A;Experimental source: strain Okra
 R;DasGupta, B.R.; Datta, A.
 Biochimie 70, 811-817, 1988
 A;Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity with
 A;Reference number: S07155; MUID:89000987; PMID:3139097
 A;Accession: S07155
 A;Molecule type: protein
 A;Residues: 2-29,'M',31-45 <DAS>
 A;Accession: S08562
 A;Molecule type: protein
 A;Residues: 442-463,'R',465-467 <DA2>
 R;Schmidt, J.J.; Sathyanarayanan, V.; DasGupta, B.R.
 Arch. Biochem. Biophys. 238, 544-548, 1985
 A;Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
 A;Reference number: S07128; MUID:85197963; PMID:3888113
 A;Accession: S07128
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 2-16 <SCH1>
 A;Accession: S08573
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 2-17 <SCH2>
 A;Accession: S08574
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 442-459 <SCH3>
 R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B.
 Nature 359, 832-835, 1992
 A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteoly
 A;Reference number: S27125; MUID:93063293; PMID:1331807
 A;Contents: annotation
 C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic syna
 C;Genetics:
 A;Gene: bont/b
 C;Function:
 A;Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2
 C;Superfamily: tetanus toxin
 C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
 F;2-441/Product: bontoxilysin B light chain #status experimental <LGH>
 F;442-1291/Product: bontoxilysin B heavy chain #status experimental <H>
 F;230,234/Binding site: zinc (His) #status predicted
 F;231/Active site: Glu #status predicted

Query Match 31.2%; Score 58; DB 1; Length 1291;
 Best Local Similarity 64.3%; Pred. No. 8.1; Mismatches 4; Indels 1; Gaps 0;
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
 Db 923 FLDFSVFWIRPK 936

RESULT 10
 I40631
 non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum
 C;Species: Clostridium botulinum
 C;Date: 12-Aug-1996 #sequence revision 12-Aug-1996 #text change 16-Jul-1999
 C;Accession: I40631; S48104; S36015
 R;Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.
 Curr. Microbiol. 28, 101-110, 1994
 A;Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botulin
 A;Reference number: I40631; MUID:94122659; PMID:7764370
 A;Accession: I40631
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA

A;Residues: 1-1291 <RES>
A;Cross-references: EMBL:X71343; NID:G296148; PIDN:CAA50482.1; PID:G296149
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinum neurotoxin gene and specific id
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48103
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 634-761, 'E', 763-841, 'M', 843, 'T', 845, 'N', 847-994 <CAM1>
A;Cross-references: EMBL:X70814; NID:G407778; PIDN:CAA50145.1; PID:G407779
A;Experimental source: non-protocytic strain 2129B (Scott)
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
A;Accession: S48104
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 634-843, 'T', 845, 'N', 847-994 <CAM2>
A;Cross-references: EMBL:X70819; NID:G407780; PIDN:CAA50150.1; PID:G407781
A;Experimental source: non-protocytic strain Eklund 2B (Colworth 229)
A;Comment: Botulinum neurotoxin type B in these strains may possess a capable catalytic s
C;Genetics:
A;Gene: Sont/b
C;Superfamily: tetanus toxin
C;Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc
F:2-441/Product: botulinum neurotoxin type B light chain #status predicted <LIGHT>
F:442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HVT>
F:230-234/Binding site: zinc (His) #status predicted
F:231/Active site: Glu #status predicted

Query Match 31.2%; Score 58; DB 2; Length 1291;
Best Local Similarity 64.3%; Pred. No. 8.1;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSPWLRVPK 14
DB 923 FLDFSVFWIRPK 936
| : : : : : |

RESULT 11
RHMSG
gonadoliberin precursor - mouse
N;Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing
N;Contains: gonadoliberin; gonadoliberin-associated protein (GAP)
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
R;Nason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolic,
Science 234, 1366-1371, 1986
A;Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible fo
A;Reference number: A47578; MUID:87069928; PMID:3024317
A;Accession: A47578
A;Molecule type: DNA
A;Residues: 1-90 <MAS>
A;Cross-references: EMBL:M14872; NID:G193576; PIDN:AAA37717.1; PID:G387175
C;Genetics:
A;Introns: 45/3; 77/3
C;Function:
A;Description: gonadoliberin stimulates pituitary secretion of lutropin and follitropin
A;Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:22-31/Product: gonadoliberin #status predicted <GLB>
F:35-90/Product: gonadoliberin-associated protein #status predicted <GAP>
F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 30.9%; Score 57.5; DB 1; Length 90;
Best Local Similarity 80.0%; Pred. No. 0.51;
Matches 12; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 20 LEG-PSLHWSYGLRP 33
| | | | |

Db 16 LEGCSSQHWYGLRP 30
| | | | |

RESULT 12
T36589
probable transmembrane protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C;Accession: T36589
R;Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z21575
A;Accession: T36589
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-502 <OLI>
A;Cross-references: EMBL:AL049826; PIDN:CAB42730.1; GSPDB:GN00070; SCOEDB:SCH24.24
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCH24.24
C;Superfamily: Mycobacterium tuberculosis probable membrane protein Rv0051

Query Match 30.6%; Score 57; DB 2; Length 502;
Best Local Similarity 50.0%; Pred. No. 4;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 18 SHLEGPSSLHWSYGLRP 33
| | | | |

Db 483 AHFEGPQVHWGKGRQP 498
| | | | |

RESULT 13
S48106
neurotoxin type E - Clostridium botulinum (fragment)
C;Species: Clostridium botulinum
C;Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C;Accession: S48106
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinum neurotoxin gene and specific
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48106
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-367 <CAM>
A;Cross-references: EMBL:X70818; NID:G407784; PIDN:CAA50149.1; PID:G407785
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

Query Match 30.4%; Score 56.5; DB 2; Length 367;
Best Local Similarity 22.4%; Pred. No. 3.3;
Matches 15; Conservative 8; Mismatches 7; Indels 37; Gaps 2;

QY 1 FNNFTVSPWLRVP 23
| | | | |

Db 297 YKNFSISFWIRPNYDNKIWNVNTIINCEDNNSGKVSINHNHNIITLQDNAGINQ 356
| | | | |

QY 24 SLHWSYG 30
| : : |

Db 357 KLAFTNYG 363
| : : |

RESULT 14
JH0256
botulinum neurotoxin type E precursor - Clostridium butyricum
C;Species: Clostridium butyricum
C;Date: 30-Jun-1992 #sequence_revision 15-May-1998 #text_change 16-Jul-1999
C;Accession: JH0256; S16145
R;Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A;Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum ty
A;Reference number: JH0256; MUID:92181428; PMID:1543481

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 6.2179 Seconds

(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-10

Perfect score: 186

Sequence: 1 FNNFTVFWLRVPKVSASHLEGPSLHMSYGLRPX 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query % | Length | DB ID | Description |
|------------|-------|---------|--------|-------|-------------|
| 1 | 112 | 60.2 | 1314 | 1 | TETX_CLOTE |
| 2 | 63 | 33.9 | 431 | 1 | PURA_LEGPN |
| 3 | 61.5 | 33.1 | 92 | 1 | GONI_RAT |
| 4 | 61 | 32.8 | 1274 | 1 | BXF_CLOBO |
| 5 | 61 | 32.8 | 1296 | 1 | BXG_CLOBO |
| 6 | 59 | 31.7 | 1295 | 1 | BXA1_CLOBO |
| 7 | 58.5 | 31.5 | 1250 | 1 | BXE_CLOBO |
| 8 | 58 | 31.2 | 1230 | 1 | BXB_CLOBO |
| 9 | 57.5 | 30.9 | 90 | 1 | GONI_MOUSE |
| 10 | 57 | 30.6 | 1051 | 1 | VP2_AHSV6 |
| 11 | 56.5 | 30.4 | 90 | 1 | GONI_RANCA |
| 12 | 56.5 | 30.4 | 1250 | 1 | BXA2_CLOBO |
| 13 | 56 | 30.1 | 1295 | 1 | BXA3_CLOBO |
| 14 | 54.5 | 29.3 | 92 | 1 | GONI_HUMAN |
| 15 | 54.5 | 29.3 | 1290 | 1 | BXC1_CLOBO |
| 16 | 54 | 29.0 | 67 | 1 | GONI_MACMU |
| 17 | 54 | 29.0 | 91 | 1 | GONI_PIG |
| 18 | 53.5 | 28.8 | 760 | 1 | AMY_CLOAB |
| 19 | 53.5 | 28.8 | 3122 | 1 | DPOZ_MOUSE |
| 20 | 52 | 28.0 | 61 | 1 | GONI_SHEEP |
| 21 | 52 | 28.0 | 63 | 1 | GONI_MESAU |
| 22 | 52 | 28.0 | 89 | 1 | GONI_XENLA |
| 23 | 52 | 28.0 | 92 | 1 | GONI_TURGE |
| 24 | 52 | 28.0 | 449 | 1 | VNSS_INSVN |
| 25 | 52 | 28.0 | 464 | 1 | VNSS_TSWV1 |
| 26 | 52 | 28.0 | 467 | 1 | VNSS_TSWVL |
| 27 | 51.5 | 27.7 | 90 | 1 | GON8_RANDY |
| 28 | 51.5 | 27.7 | 91 | 1 | GON1_ORYLA |
| 29 | 51.5 | 27.7 | 521 | 1 | UBP3_HUMAN |
| 30 | 51 | 27.4 | 95 | 1 | GONI_MORSA |
| 31 | 51 | 27.4 | 265 | 1 | DDHC_RHOSU |
| 32 | 51 | 27.4 | 292 | 1 | CC22_ORYSA |
| 33 | 51 | 27.4 | 1196 | 1 | BXCN_CLOBO |

| | | | | | |
|----|------|------|------|---|------------|
| 34 | 50.5 | 27.2 | 537 | 1 | COX1_SCHPO |
| 35 | 50.5 | 27.2 | 1450 | 1 | SREJ_STRPU |
| 36 | 50 | 26.9 | 241 | 1 | CRTA_RHOCA |
| 37 | 50 | 26.9 | 842 | 1 | ABCG_HUMAN |
| 38 | 49.5 | 26.6 | 95 | 1 | GONI_PAGMA |
| 39 | 49.5 | 26.6 | 95 | 1 | GONI_SPAAU |
| 40 | 49.5 | 26.6 | 3130 | 1 | DPOZ_HUMAN |
| 41 | 49 | 26.3 | 523 | 1 | C9BI_GLYEC |
| 42 | 49 | 26.3 | 551 | 1 | SYK_BRUSE |
| 43 | 49 | 26.3 | 551 | 1 | SYK_BRUSE |
| 44 | 48.5 | 26.1 | 92 | 1 | GONI_CAVPO |
| 45 | 48.5 | 26.1 | 5217 | 1 | HTS1_COCCA |

ALIGNMENTS

| | |
|------------|--|
| RESULT 1 | |
| TETX_CLOTE | |
| ID | TETX_CLOTE STANDARD; PRT; 1314 AA. |
| AC | P04958; |
| DT | 13-AUG-1987 (Rel. 05, Created) |
| DT | 13-AUG-1987 (Rel. 05, last sequence update) |
| DT | 10-OCT-2003 (Rel. 42, last annotation update) |
| DE | Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin) [Contains: |
| DE | Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy |
| DE | chain (Tetanus toxin chain H)]. |
| GN | TETX OR CTP60. |
| OS | Clostridium tetani. |
| OC | Plasmod p88, and Plasmid 75 Kbp. |
| OC | Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; |
| OC | Clostridium. |
| OX | NCBI_TaxID=1513; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | PLASMOD=75 Kbp; |
| RC | MEDLINE=87053814; PubMed=3536478; |
| RA | Eisel U., Jarasch W., Goretzki K., Henschen A., Engels J., |
| RA | Weller U., Hudel M., Habermann E., Niemann H.; |
| RT | "Tetanus toxin: primary structure, expression in E. coli, and |
| RT | homology with botulinum toxins."; |
| RL | EMBO J. 5:2495-2502(1986). |
| RN | [2] |
| RP | SEQUENCE FROM N.A. |
| RC | STRAIN=CN3911; PLASMOD=75 Kbp; |
| RC | MEDLINE=87040747; PubMed=3774547; |
| RA | Fairweather N.F., Lyness V.A.; |
| RT | "The complete nucleotide sequence of tetanus toxin."; |
| RL | Nucleic Acids Res. 14:7809-7812(1986). |
| RN | [3] |
| RP | SEQUENCE FROM N.A. |
| RC | STRAIN=Massachusetts / E88; PLASMOD=p88; |
| RC | MEDLINE=22457253; PubMed=12552129; |
| RA | Brueggemann H., Baeumer S., Fricke W.F., Wierzer A., Liesegang H., |
| RA | Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A., |
| RT | Gotteschalk G.; |
| RT | "The genome sequence of Clostridium tetani, the causative agent of |
| RT | tetanus disease."; |
| RL | Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003). |
| RN | [4] |
| RP | SEQUENCE OF 742-1314 FROM N.A. |
| RC | PLASMOD=75 Kbp; |
| RC | MEDLINE=86085672; PubMed=3510187; |
| RA | Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.; |
| RT | "Cloning, nucleotide sequencing, and expression of tetanus toxin |
| RT | fragment C in Escherichia coli."; |
| RL | J. Bacteriol. 165:21-27(1986). |
| RN | [5] |
| RP | PARTIAL SEQUENCE, AND DISULFIDE BONDS. |
| RC | MEDLINE=90201034; PubMed=2108021; |
| RA | Kriegelstein K., Henschen A., Weller U., Habermann E.; |
| RT | "Arrangement of disulfide bridges and positions of sulfhydryl groups |
| RT | in tetanus toxin."; |

| | |
|--------|-------------|
| P07657 | schizosacch |
| Q26627 | strongyloce |
| P17055 | rhodobacter |
| Q9np58 | homo sapien |
| P70074 | pagrus majo |
| P51319 | sparus auro |
| O60673 | homo sapien |
| Q8ycm8 | glycyrhiza |
| P59225 | brucella me |
| O54713 | cavia porce |
| Q01886 | cochliobolu |

RL Eur. J. Biochem. 188:39-45 (1990).
RN [6]
RX PARTIAL SEQUENCE.
RA MEDLINE=92037649; PubMed=1935979;
RZ Krieglstein K.G., Heschel A.H., Weller U., Habermann E.;
RT "Limited proteolysis of tetanus toxin. Relation to activity and
RT identification of cleavage sites";
RL Eur. J. Biochem. 202:41-51 (1991).
RN [7]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93010948; PubMed=1396558;
RA Schiavo G., Poullain B., Rossetto O., Benfenati F., Tauc L.,
RA Montecucco C.;
RT "Tetanus toxin is a zinc protein and its inhibition of
RT neurotransmitter release and protease activity depend on zinc";
RL EMBO J. 11:3577-3583 (1992).
RN [8]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=93063293; PubMed=1331807;
RA Schiavo G., Benfenati F., Poullain B., Rossetto O., de Laureto P.P.,
RA Dasgupta B.R., Montecucco C.;
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin";
RL Nature 359:832-835 (1992).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
RX MEDLINE=97475217; PubMed=9334741;
RA Unland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
RA Sax M.;
RT "Structure of the receptor binding fragment HC of tetanus
RT neurotoxin";
RL Nat. Struct. Biol. 4:788-792 (1997).
CC -|- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
CC BOND OF SYNAPTOSOMAL-2.
CC -|- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-|-Phe-77 bond in
CC synaptobrevin 2.
CC -|- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -|- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
CC AND ARE NON-TOXIC AFTER SEPARATION.
CC -|- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
CC GANGLIOSIDE RECEPTORS.
CC -|- SIMILARITY: Belongs to peptidase family M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X04436; CAA28033.1; -
CC EMBL; X06214; CAA29564.1; -
CC EMBL; AF528097; AAC37454.1; -
CC EMBL; M12739; AAA23282.1; -
CC PIR; A25689; BTCLTN.
CC PDB; 1AF9; 29-APR-98.
CC PDB; 1ABD; 14-OCT-98.
CC PDB; 1DOH; 27-MAR-00.
CC PDB; 1DFQ; 24-MAR-00.
CC PDB; 1DIW; 24-MAR-00.
CC PDB; 1DLL; 24-MAR-00.
CC PDB; 1FV3; 05-SEP-01.
CC MEROPS; M27.001; -
CC InterPro; IPR006985; ConA like lec_gl.
CC InterPro; IPR002160; Kunitz_legume.
CC InterPro; IPR006025; Pept_M_Zn_BS.

DR InterPro; IPR000395; Peptidase_M27.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PD00760; BONTOKILYSIN.
DR PRODOM; PD001963; Bontokilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
KW 3D-structure; Complete proteome.
FT INIT MET 0
FT CHAIN 0
FT CHAIN 1 456 TETANUS TOXIN LIGHT CHAIN.
FT CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.
FT METAL 232 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 233 BY SIMILARITY.
FT METAL 236 ZINC (CATALYTIC) (BY SIMILARITY).
FT TRANSMEM 226 246 POTENTIAL.
FT TRANSMEM 669 689 POTENTIAL.
FT DISULFID 438 466 INTERCHAIN.
FT DISULFID 1076 1092
FT HELIX 876 882
FT TURN 883 883
FT STRAND 884 891
FT TURN 892 893
FT STRAND 894 897
FT STRAND 904 907
FT TURN 909 910
FT STRAND 912 915
FT STRAND 920 925
FT TURN 928 929
FT STRAND 932 935
FT HELIX 938 940
FT TURN 941 946
FT STRAND 949 956
FT HELIX 962 968
FT TURN 969 970
FT STRAND 972 977
FT STRAND 980 981
FT HELIX 983 985
FT STRAND 987 995
FT TURN 996 997
FT STRAND 998 1004
FT TURN 1006 1007
FT STRAND 1010 1016
FT STRAND 1020 1020
FT TURN 1021 1022
FT STRAND 1031 1037
FT TURN 1039 1040
FT STRAND 1042 1047
FT TURN 1048 1049
FT STRAND 1050 1056
FT TURN 1058 1059
FT STRAND 1068 1074
FT TURN 1079 1080
FT STRAND 1082 1091
FT HELIX 1097 1105
FT TURN 1106 1107
FT STRAND 1112 1112
FT STRAND 1114 1114
FT TURN 1116 1117
FT STRAND 1120 1120
FT STRAND 1122 1122
FT TURN 1123 1124
FT STRAND 1127 1131
FT HELIX 1132 1134
FT TURN 1135 1136
FT STRAND 1137 1141
FT TURN 1144 1145
FT STRAND 1148 1152
FT STRAND 1155 1158
FT TURN 1159 1162
FT STRAND 1163 1166
FT TURN 1173 1178
FT STRAND 1184 1185
FT STRAND 1188 1188
FT STRAND 1190 1190

Query Match 60.2%; Score 112; DB 1; Length 1314;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVFWLRVVKVASHLE 21
Db 946 FNNFTVFWLRVVKVASHLE 966

RESULT 2

PURA LEGPN STANDARD; PRT; 431 AA.
AC QBRN2; 2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase)
DE (AdSS) (AMPase).
GN PURA.
OS Legionella pneumophila.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Legionellaceae; Legionella.
OX NCBI_TaxID=446;
RN [1]
RP SEQUENCE FROM N.A.
RA Rankin S., Li Z., Isberg R.R.;
RT "Macrophage induced genes of Legionella pneumophila: protection from
reactive intermediates and solute imbalance during intracellular
growth";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Plays an important role in the de novo pathway of purine
nucleotide biosynthesis.
CC -1- CATALYTIC ACTIVITY: GTP + IMP + L-aspartate = GDP + phosphate +
adenylosuccinate.
CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -1- PATHWAY: AMP biosynthesis; first committed step.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the adenylosuccinate synthetase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL; AF480918; AAM00648.1; -
DR HAMAP; MF_00011; -; 1.
DR InterPro; IPR001114; Aseucc synthetase.
DR Pfam; PF00709; Adenylosucc synt; 1.
DR ProDom; PD001188; Aseucc synthetase; 1.
DR TIGRFAMs; TIGR00184; PURA; 1.
DR PROSITE; PS01265; ADENYLOSUCIN SYN 1; 1.
DR Purine biosynthesis; ligase; GTP (POTENTIAL).
KW NP_BIND 13 19
FT ACT_SITE 141 141 BY SIMILARITY.
FT ACT_SITE 148 148 BY SIMILARITY.
FT METAL 14 14 MAGNESIUM (BY SIMILARITY).
FT METAL 41 41 MAGNESIUM (VIA CARBONYL OXYGEN) (BY
SIMILARITY)
FT SEQUENCE 431 AA; 47381 MW; D446C19A8D5S0774 CRC64;

Query Match 33.9%; Score 63; DB 1; Length 431;
Best Local Similarity 36.4%; Pred. No. 0.25;
Matches 12; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Qy 1 FNNFTVFWLRVVKVASHLEGPFLWSVGLRP 33
Db 168 YHNFVLTQYKQFVNDLESLLGESLOWARELRP 200

RESULT 3

GONL RAT STANDARD; PRT; 92 AA.
AC P07450;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberin I precursor (Contains: Gonadoliberin I (LH-RH I)
(luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
I).
GN GNRH1 OR GNRH;
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
of gonadotropin-releasing hormone and prolactin release-inhibiting
factor in human and rat";
RT factor in human and rat";
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89384661; PubMed=2475669;
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT "The rat gonadotropin-releasing hormone: SH locus: structure and
hypothalamic expression";
RT Mol. Endocrinol. 3:1257-1262(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93105480; PubMed=1468115;
RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic
luteinizing hormone-releasing hormone mRNA";
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN [4]
RP SEQUENCE OF 1-47 FROM N.A.
RX MEDLINE=87149087; PubMed=3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
DNA locus";
RL Science 235:1514-1517(1987).
CC -1- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
the secretion of both luteinizing and follicle-stimulating
hormones.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Central nervous system.
CC -1- SIMILARITY: Belongs to the GNRH family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL; S50870; AAB24572.1; -
DR EMBL; M12579; AAA41263.1; -
DR EMBL; M1670; AAA41264.1; -
DR EMBL; M15527; AAA42141.1; ALT_SEQ.
DR EMBL; M15529; AAA42139.1; -
DR EMBL; M15528; -; NOT_ANNOTATED_CDS.
DR PIR; A40147; RHRTG.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; GonadoliberinI.

PFAM: PF00446; GNRH; 1.
DR PRINTS; PRO1541; GONADOLIBERNI.
DR PROSITE; PS00473; GNRH; 1.
KW cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
Placenta; Signal; Pyroglutamate carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;
Query Match 33.1%; Score 61.5; DB 1; Length 92;
Best Local Similarity 51.7%; Pred. No. 0.076; 4; Indels 7; Gaps 2;
Matches 15; Conservative 3; Mismatches 4; Indels 7; Gaps 2;
QY 12 VPKVAS-----HLRG-PSLHWSYGLRP 33
Db 4 IPKLMAAVLLTVCLGSCSQHWSYGLRP 32
RESULT 4
EXP_CLOBO STANDARD; PRT; 1274 AA.
ID AC P30596;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BoNT/F)
DE (Bontoxilysin F).
DE BOTF.
GN Clostridium botulinum.
OS Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23387;
RX MEDLINE=93012902; PubMed=1398040;
RA East A.K., Richardson P.T., Allaway D., Collins M.D.,
RA Roberts T.A., Thompson D.E.;
RT "Sequence of the gene encoding type F neurotoxin of Clostridium
botulinum";
RL FEWS Microbiol. Lett. 75:225-230(1992).
RN [2]
RP SEQUENCE OF 1-64 FROM N.A.
RC STRAIN=Hobbs FT10; PubMed=7764998;
RX MEDLINE=94297488; PubMed=8408542;
RA East A.K., Collins M.D.;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinum neurotoxin gene and
nontoxic component in nonproteolytic Clostridium botulinum type F.";
RL Curr. Microbiol. 29:69-77(1994).
RN [3]
RP SEQUENCE OF 634-1002 FROM N.A.
RX MEDLINE=94230352; PubMed=8175689;
RA Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
RA Roques B., Fykse E.M., Suedhof T.C., Jahn R., Niemann H.;
RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
F botulinum neurotoxins and tetanus toxin";
RL J. Biol. Chem. 269:12764-12772(1994).
CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED

AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-|-LYS-59
BOND OF SYNAPTOBREVIN-1 AND -2.
CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
detected action on small molecule substrates.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
heavy chain (H). The light chain has the pharmacological activity,
while the N- and C-terminal of the heavy chain mediate channel
formation and toxin binding, respectively.
CC -!- MISCELLANEOUS: There are seven antigenically distinct forms of
botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC -!- SIMILARITY: Belongs to peptidase family M27.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; M92906; AAA23263.1; -
DR EMBL; S73676; AAC60475.1; -
DR EMBL; X70820; CAA50151.1; -
DR EMBL; X70816; CAA50147.1; -
DR PIR; I40813; I40813.
DR PIR; I40813; I40813.
DR FIR; S48109; S48109.
DR HSRP; P10845; 3BTA.
DR MEROPS; M27.002; -
DR InterPro; IPR008985; ConA_like_lect_gl.
DR InterPro; IPR002160; Kunitz_legume.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000395; Peptidase_M27.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PRO0760; BONTOLILYSIN.
DR PRODOM; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT CHAIN 1 436
FT METAL 227 227
FT ACT_SITE 228 228
FT METAL 231 231
FT DISULFID 429 445
SQ SEQUENCE 1274 AA; 146709 MW; 5B99756A7438B921 CRC64;
Query Match 32.8%; Score 61; DB 1; Length 1274;
Best Local Similarity 57.1%; Pred. No. 1.5;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNNFTVSFWLRVPK 14
Db 930 YQNFISFWVRIPK 943
RESULT 5
EXP_CLOBO STANDARD; PRT; 1296 AA.
ID AC Q60393;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BoNT/G)
DE (Bontoxilysin G).
GN BOTG.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
CC Clostridium.

OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=113 / 30;
RC MEDLINE=94092745; PubMed=8268233;
RA Campbell K., Collins M.D., East A.K.;
RT "Nucleotide sequence of the gene coding for Clostridium botulinum
RT (Clostridium argentinense) type G neurotoxin: genealogical comparison
RT with other clostridial neurotoxins";
RL Biochim. Biophys. Acta 1216:487-491(1993).
CC -|- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC
CC -|- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC
CC -|- COFACTOR: Binds 1 zinc ion per subunit (by similarity). (L) and a
CC heavy chain (H). The light chain has the pharmacological activity,
CC while the N- and C-terminal of the heavy chain mediate channel
CC formation and toxin binding, respectively.
CC
CC -|- SUBCELLULAR LOCATION: Secreted (By similarity).
CC
CC -|- MISCELLANEOUS: There are seven antigenically distinct forms of
CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC
CC -|- SIMILARITY: Belongs to peptidase family M27.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X74162; CRA52275.1; --
CC HSP; P10845; 3BTA.
CC
CC MEMOPS; M27 002; --
CC InterPro; IPR008985; ConA_like_lsc_gl.
CC InterPro; IPR002160; Kunitz_legume.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR000395; Peptidase_M27.
CC Pfam; PF01742; Peptidase_M27; 1.
CC PRINTS; PR00760; BONTOKILYSIN.
CC ProDom; P0001963; Bontoxilysin; 1.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC Neurotoxin; Hydrolase; Metalloprotease; Zinc.
CC
CC INIT MET 0 BY SIMILARITY.
CC CHAIN 1 441 BOTULINUM NEUROTOXIN G, LIGHT-CHAIN.
FT CHAIN 442 1296 BOTULINUM NEUROTOXIN G, HEAVY-CHAIN.
FT CHAIN 229 230 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 230 230 BY SIMILARITY.
FT ACT_SITE 230 230 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 435 449 INTERCHAIN (PROBABLE).
SQ SEQUENCE 1296 AA; 149013 MW; DC8E47E15F665C31 CRC64;

Query Match 32.8%; Score 61; DB 1; Length 1296;
Best Local Similarity 38.1%; Pred. No. 1.6;
Matches 8; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRPKVSASHLE 21
Db 929 FDNFSINFWRTPKYNNDIQ 949

RESULT 6
EXAL_CLOBO STANDARD; PRT; 1295 AA.
ID BXA_CLOBO
AC P10845; P01561; P18639;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)
DE (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
DE chain; Botulinum neurotoxin A, heavy-chain].
DE BOTA OR BNA OR ATX.
OS Clostridium botulinum.
CC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
CC Clostridium botulinum.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 2916;
RX MEDLINE=90235864; PubMed=2185020;
RA Thompson D.E., Brehm J.K., Oultram J.D., Swinfield T.-J.,
RA Shone C.C., Atkinson T., Melling J., Minton N.P.;
RT "The complete amino acid sequence of the Clostridium botulinum type A
RT neurotoxin, deduced by nucleotide sequence analysis of the encoding
RT gene";
RL Eur. J. Biochem. 189:73-81(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=62A;
RX MEDLINE=90264400; PubMed=2160960;
RA Binz B., Kuaronzo H., Willie M., Frevant J., Wernars K., Niemann H.;
RT "The complete sequence of botulinum neurotoxin type A and comparison
RT with other clostridial neurotoxins";
RL J. Biol. Chem. 265:9153-9158(1990).
RN [3]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN=62A;
RX MEDLINE=97016917; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding
RT components of the botulinum toxin complex in proteolytic Clostridium
RT botulinum types A, B, and F: evidence of chimeric sequences in the
RT gene encoding the nontoxic nonhemagglutinin component";
RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
RN [4]
RP SEQUENCE OF 1-34 FROM N.A.
RC STRAIN=Hall;
RX MEDLINE=89350859; PubMed=2669749;
RA Betley M.J., Somers B., Dasgupta B.R.;
RT "Characterization of botulinum type A neurotoxin gene: delineation of
RT the N-terminal encoding region";
RL Biochem. Biophys. Res. Commun. 162:1388-1395(1989).
RN [5]
RP SEQUENCE OF 1-18 FROM N.A.
RC STRAIN=Type A NIH;
RX MEDLINE=96096783; PubMed=8521962;
RA Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;
RT "Molecular characterization of two forms of nontoxic-nonhemagglutinin
RT components of Clostridium botulinum type A progenitor toxins";
RL FEBS Lett. 376:41-44(1995).
RN [6]
RP SEQUENCE OF 1-16.
RX MEDLINE=84178501; PubMed=6370252;
RA Schmidt J.J., Sartymoorthy V., Dasgupta B.R.;
RT "Partial amino acid sequence of the heavy and light chains of
RT botulinum neurotoxin type A";
RL Biochem. Biophys. Res. Commun. 119:900-904(1984).
RN [7]
RP SEQUENCE OF 1-46.
RA Dasgupta B.R., Foley J., Niece R.;
RT "Partial sequence of the light chain of botulinum neurotoxin type A";
RL Biochemistry 26:4162-4162(1987).
RN [8]
RP SEQUENCE OF 1-5 AND 444-456.
RX MEDLINE=91120847; PubMed=2126206;
RA Dasgupta B.R., Dekleva M.L.;
RT "Botulinum neurotoxin type A: sequence of amino acids at the
RT N-terminus and around the nicking site";
RL Biochimie 72:661-664(1990).
RN [9]

SEQUENCE OF 448-464 AND 872-895.
 MEDLINE=9024662; PubMed=3178218;
 RT "Botulinum neurotoxin type A: Cleavage of the heavy chain into two
 RT halves and their partial sequences.";
 RL Arch. Biochem. Biophys. 266:142-151(1998).
 RN [10]
 RP SEQUENCE OF 448-482.
 RX MEDLINE=95285016; PubMed=3896784;
 RA Shone C.C., Hambleton P., Melling J.,
 RT "Inactivation of Clostridium botulinum type A neurotoxin by trypsin
 RT and purification of two tryptic fragments. Proteolytic action near
 RT the COOH-terminus of the heavy subunit destroys toxin-binding
 RT activity.";
 RL Eur. J. Biochem. 151:75-82(1985).
 RN [11]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94063091; PubMed=8243676;
 RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
 RT Benfenati F., Wilson M.C., Montecucco C.,
 RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
 RT COOH-terminal peptide bonds.";
 RL FEBS Lett. 335:99-103(1993).
 RN [12]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94124495; PubMed=8294407;
 RA Biaz T., Blas J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
 RT Jahn R., Niemann H.,
 RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
 RL J. Biol. Chem. 269:1617-1620(1994).
 RN [13]
 RP MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.
 RX MEDLINE=21556941; PubMed=11700044;
 RA Rigoni M., Caccin P., Johnson E.A., Montecucco C., Rossetto O.,
 RT "Site-directed mutagenesis identifies active-site residues of the
 RT light chain of botulinum neurotoxin type A.";
 RL Biochem. Biophys. Res. Commun. 288:1237(2001).
 RN [14]
 RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
 RX MEDLINE=98455071; PubMed=9783750;
 RA Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.,
 RT "Crystal structure of botulinum neurotoxin type A and implications
 RT for toxicity.";
 RL Nat. Struct. Biol. 5:898-902(1998).
 CC -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
 CC binds with high affinity to peripheral neuronal presynaptic
 CC membrane, is then internalized by receptor-mediated endocytosis.
 CC The C-terminus of the heavy chain (H) is responsible for the
 CC adherence of the toxin to the cell surface while the N-terminus
 CC mediates transport of the light chain from the endocytic vesicle
 CC to the cytosol. After translocation, the light chain (L)
 CC hydrolyzes the 197-Glu-|-Arg-198 bond in SNAP-25, thereby blocking
 CC neurotransmitter release. Inhibition of acetylcholine release
 CC results in flaccid paralysis, with frequent heart or respiratory
 CC failure.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit.
 CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
 CC heavy chain (H).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PHARMACEUTICAL: Available under the name BOTOX (Allergan) for
 CC the treatment of strabismus and blepharospasm associated with
 CC dystonia and cervical dystonia. Also used for the treatment of
 CC hemifacial spasm and a number of other neurological disorders
 CC characterized by abnormal muscle contraction.
 CC -!- MISCELLANEOUS: There are seven antigenically distinct forms of
 CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
 CC -!- SIMILARITY: Belongs to peptidase family M27.
 CC -!- DATABASE: NAME=BOTOX product information web site;
 CC WWW="http://www.botox.com/index.jsp?hp&productinfo".
 CC -!- DATABASE: NAME=Protein Spotlight;

NOTE=Issue 19 of February 2002;
 WWW="http://www.expasy.org/spotlight/articles/spl019.html".
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR ENBL; X52086; CAA36289.1; -;
 DR ENBL; M30196; AAA3282.1; -;
 DR ENBL; X92973; CAA63551.1; -;
 DR ENBL; D67030; BAA1051.1; -;
 DR ENBL; M27892; AAA23269.1; -;
 DR PIR; A35294; BTCLAB.
 DR PDB; 3BTA; 01-002-99.
 DR MEROPS; M27_002; -;
 DR InterPro; IPR008985; ConA like lec_gl.
 DR InterPro; IPR002160; Kunitz_legume.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR000395; Peptidase_M27.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC PROTEASE; 1.
 DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
 KW Pharmaceutical; 3D-structure.
 FT INIT_MET 0 0
 FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
 FT CHAIN 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
 FT METAL 222 222 ZINC (CATALYTIC).
 FT ACT_SITE 223 223 ZINC (CATALYTIC).
 FT METAL 226 226 ZINC (CATALYTIC).
 FT METAL 261 261 ZINC (CATALYTIC).
 FT DISULFID 429 453 INTERCHAIN.
 FT DISULFID 1234 1279 POTENTIAL.
 FT TRANSMEM 656 646 POTENTIAL.
 FT TRANSMEM 655 675 POTENTIAL.
 FT VARIANT 26 26 V -> A.
 FT MUTAGEN 261 261 E->A: DRASTIC DECREASE IN ENZYMATIC
 FT ACTIVITY.
 FT MUTAGEN 265 265 F->A: DECREASES ENZYMATIC ACTIVITY.
 FT MUTAGEN 365 365 Y->A: DECREASES ENZYMATIC ACTIVITY.
 FT CONFLICT 1 1 P -> Q (IN REF. 1).
 FT CONFLICT 479 479 E -> P (IN REF. 9).
 FT CONFLICT 875 875 T -> L (IN REF. 8).
 FT CONFLICT 891 891 S -> K (IN REF. 8).
 SQ SEQUENCE 1295 AA; 149322 MW; 858342F754862579 CRC64;
 Query Match 31.7%; Score 59; DB 1; Length 1295;
 Best Local Similarity 57.1%; Pred. No. 3;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 FNNFTVSFWLRVVK 14
 DB 937 YENFSTFWIRPK 950
 RESULT 7
 BXE_CLOSO
 ID_BXE_CLOSO STANDARD; PRT: 1250 AA.
 AC Q00496;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BoNT/E)
 DE (Bontoxilysin E).
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OC NCBI_taxid=1491;

RN RP SEQUENCE FROM N.A.
 RC STRAIN=Beluga;
 RX MEDLINE=92181428; PubMed=1543481;
 RA "Sequences of the botulinum neurotoxin E derived from Clostridium
 RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
 RT ATCC 43181 and ATCC 43755).";
 RL Biochem. Biophys. Res. Commun. 183:107-113 (1992).
 [2]
 RN RP SEQUENCE FROM N.A.
 RP MEDLINE=92174922; PubMed=1541280;
 RX Whelan S.M., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;
 RA "The complete amino acid sequence of the Clostridium botulinum type-E
 RT neurotoxin, derived by nucleotide-sequence analysis of the encoding
 RT gene.";
 RL Eur. J. Biochem. 204:657-667 (1992).
 [3]
 RN RP SEQUENCE OF 1-251 FROM N.A.
 RP MEDLINE=90264400; PubMed=2160960;
 RX Binz T., Kurazono H., Wille M., Frevert J., Wernars K., Niemann H.;
 RA "The complete sequence of botulinum neurotoxin type A and comparison
 RT with other clostridial neurotoxins.";
 RL J. Biol. Chem. 265:9153-9158 (1990).
 [4]
 RN RP SEQUENCE OF 1-13.
 RP MEDLINE=85197963; PubMed=3888113;
 RX Schmidt J.J., Sathyanarayanan V., Dasgupta B.R.;
 RA "Partial amino acid sequences of botulinum neurotoxins types B and
 RT E.";
 RL Arch. Biochem. Biophys. 238:544-548 (1985).
 [5]
 RN RP SEQUENCE OF 419-426.
 RP MEDLINE=90344918; PubMed=2116911;
 RX Gimenez J.A., Dasgupta B.R.;
 RA "Botulinum neurotoxin type E fragmented with endoprotease Lys-C
 RT reveals the site trypsin nicks and homology with tetanus
 RT neurotoxin.";
 RL Biochimie 72:213-217 (1990).
 [6]
 RN RP IDENTIFICATION OF SUBSTRATE.
 RP MEDLINE=94063091; PubMed=8243676;
 RX Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
 RA Benfenati F., Wilson M.C., Montecucco C.;
 RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
 RT COOH-terminal peptide bonds.";
 RL FEBS Lett. 335:99-103 (1993).
 [7]
 RN RP IDENTIFICATION OF SUBSTRATE.
 RP MEDLINE=94124495; PubMed=8294407;
 RX Binz T., Biasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
 RA Jahn R., Niemann H.;
 RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
 RL J. Biol. Chem. 269:1617-1620 (1994).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-|-ILE-
 CC 181 BOND IN SNAP-25.
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (by similarity). (L) and a
 CC heavy chain (H). The light chain has the pharmacological activity,
 CC while the N- and C-terminal of the heavy chain mediate channel
 CC formation and toxin binding, respectively.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of
 CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
 CC -1- SIMILARITY: Belongs to peptidase family M27.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X62089; CAA43999.1; -;
 DR EMBL; X62683; CAA44558.1; -;
 DR FIR; S08575; S08575.
 DR PIR; S21178; S21178.
 DR HSSP; P10845; 3BTA.
 DR MEROPS; M27.002; -;
 DR InterPro; IPR008985; ConA_like_lec_g1.
 DR InterPro; IPR002160; Kunitz_legume.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR000395; Peptidase_M27.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 RN INIT_MET 0
 FT CHAIN 1 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
 FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 212 212 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 411 425 INTERCHAIN (PROBABLE).
 FT CONFLICT 176 176 R -> G (IN REF. 2).
 FT CONFLICT 197 197 C -> S (IN REF. 2 AND 3).
 FT CONFLICT 339 339 R -> A (IN REF. 2).
 FT CONFLICT 772 772 I -> L (IN REF. 2).
 FT CONFLICT 962 963 FE -> LQ (IN REF. 2).
 FT CONFLICT 966 966 R -> A (IN REF. 2).
 FT CONFLICT 1194 1194 N -> NN (IN REF. 2).
 FT CONFLICT 1250 1250 N -> NN (IN REF. 2).
 SQ SEQUENCE 1250 AA; 143712 MW; D9FC26DDA041E34 CRC64;
 Query Match 31.5%; Score 58.5; DB 1; Length 1250;
 Best Local Similarity 23.6%; Pred. No. 3.4;
 Matches 13; Conservative 9; Mismatches 4; Indels 29; Gaps 2;
 Qy 1 FNNFTVSFWLRVP-----KVASHLEGPSTLWSY 29
 Db 911 YKNFSISFWVRIPENYDNKIVNNVEYTIINCVRDNNNGKVSLEHNE---ILWTF 962
 RESULT 8
 EXB_CLOSO STANDARD; PRT; 1290 AA.
 ID AC P10844; P10843;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BoNT/B)
 DE (Bontoxilysin B).
 GN BOTB.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92384550; PubMed=1514783;
 RA Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
 RA Minton N.P.;
 RT "Molecular cloning of the Clostridium botulinum structural gene
 RT encoding the type B neurotoxin and determination of its entire
 RT nucleotide sequence.";
 RL Appl. Environ. Microbiol. 58:2345-2354 (1992).
 [2]

SEQUENCE OF 35-245 FROM N.A.
 STRAIN=NCTC 7273;
 Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
 Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 [3]
 RY SEQUENCE OF 633-993 FROM N.A.
 STRAIN=NCTC 7273;
 MEDLINE=8408542;
 Campbell K., East A.K., Collins M.D.;
 "Gene probes for identification of the botulinum neurotoxin gene and
 specific identification of neurotoxin types B, E, and F";
 J. Clin. Microbiol. 31:2255-2262(1993).
 [4]
 RN SEQUENCE OF 1-44 AND 441-466.
 STRAIN=OKRA;
 MEDLINE=85197963; PubMed=3889113;
 Schmidt J.J., Sathiyamoorthy V., Dasgupta B.R.;
 "Partial amino acid sequences of botulinum neurotoxins types B and
 E";
 Arch. Biochem. Biophys. 238:544-548(1985).
 [6]
 RN IDENTIFICATION AS ZINC-PROTEASE.
 MEDLINE=93054694; PubMed=1423690;
 Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;
 "Botulinum neurotoxins are zinc proteins";
 J. Biol. Chem. 267:23479-23483(1992).
 [7]
 RN IDENTIFICATION OF SUBSTRATE.
 MEDLINE=93063293; PubMed=1331807;
 Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
 Dasgupta B.R., Montecucco C.;
 "Tetanus and botulinum-B neurotoxins block neurotransmitter release
 by proteolytic cleavage of synaptobrevin";
 Nature 359:832-835(1992).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-|-PHE-77 BOND OF
 SYNAPTOSOMAL-2.
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 detected action on small molecule substrates.
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
 heavy chain (H). The light chain has the pharmacological activity,
 while the N- and C-terminal of the heavy chain mediate channel
 formation and toxin binding, respectively.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of
 botulinum neurotoxin: Types A, B, C, D, E, F, and G.
 CC -1- SIMILARITY: Belongs to peptidase family M27.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M81186; AAA23211.1; -
 CC EMBL; Z11934; CAA77991.1; -
 CC EMBL; X70817; CAA50148.1; -

PIR; A48940; A48940.
 PDB; 1EPW; 01-NOV-00.
 DR PDB; 1F31; 01-NOV-00.
 DR PDB; 1F82; 16-AUG-00.
 DR PDB; 1F83; 16-AUG-00.
 DR PDB; 1FQ4; 06-DEC-00.
 DR PDB; 1G9A; 13-NOV-02.
 DR PDB; 1G9B; 13-NOV-02.
 DR PDB; 1G9C; 13-NOV-02.
 DR PDB; 1G9D; 13-NOV-02.
 DR PDB; 1I1E; 21-NOV-01.
 DR MEROPS; M27.002; -
 DR InterPro; IPR008985; ConA_like_lec_gl.
 DR InterPro; IPR002160; Kunitz_legume.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR000395; Peptidase_M27.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOTOXILYSIN.
 DR PRODOM; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
 3D-structure.
 FT INIT_MET 0
 FT CHAIN 1 440
 FT CHAIN 441 1290
 FT METAL 229 229
 FT ACT_SITE 230 230
 FT METAL 233 233
 FT DISULFID 436 445
 FT CONFLICT 29 29
 FT CONFLICT 217 217
 FT CONFLICT 224 224
 FT CONFLICT 463 463
 FT SEQUENCE 1290 AA; 150670 MW; D21746E2C024DF43 CRC64;
 SQ
 Query Match 31.2%; Score 58; DB 1; Length 1290;
 Best Local Similarity 64.3%; Pred. No. 4.1; Indels 0; Gaps 0;
 Matches 9; Conservative 4; Mismatches 1;
 QY 1 FNNFTVSEFWLRVVK 14
 Db 922 PLDFSVSWLRPK 935
 RESULT 9
 GONI_MOUSE STANDARD; PRT; 90 AA.
 ID GONI_MOUSE STANDARD; PRT; 90 AA.
 AC P13562;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)
 DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 DE hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
 DE I].
 GN GNRIH OR GNRIH.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 OX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87069828; PubMed=3024317;
 RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
 RA Phillips H.S., Nikolics K., Seeburg P.H.;
 RT "A deletion truncating the gonadotropin-releasing hormone gene is
 RT responsible for hypogonadism in the hpg mouse";
 RL Science 234:1366-1371(1986).
 CC -1- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the GnRH family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14872; AAA37717.1; -;
DR PIR; A47578; RHMSG.
DR MGD; MGI-95789; GnRH.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; Gonadoliberin1.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PRO1541; GONADOLIBERN1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 21 PROGNADOLIBERIN I.
FT CHAIN 22 90 GONADOLIBERIN I.
FT PEPTIDE 22 31 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT ACT_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
FT SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;
SQ
Query Match 30.9%; Score 57.5; DB 1; Length 90;
Best Local Similarity 80.0%; Pred. No. 0.27;
Matches 12; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 20 LEG-PSLHWSYGLRP 33
DB 16 LEGCSSQHWYGLRP 30
RESULT 10
ID VP2_AHSV6 STANDARD; PRT; 1051 AA.
AC 071024;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Outer capsid protein VP2.
GN S2 OR L2.
OS African horse sickness virus 6 (AHSV-6) (African horse sickness virus
OS (serotype 6)).
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX NCBI_TaxID=86060;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98278331; PubMed=9617769;
RA Williams C.F., Inoue T., Lucas A.-M., Zanotto P., Roy P.,
RT "The complete sequence of four major structural proteins of African
RT horse sickness virus serotype 6: evolutionary relationships within
RT and between the orbiviruses.";
RL Virus Res. 53:53-73(1998).
CC -!- FUNCTION: THE VP2 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VPS)
CC -!- WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID. IT IS THE
CC MAJOR TARGET OF THE HOST IMMUNOGENIC RESPONSE.
CC -!- SIMILARITY: Belongs to the reoviruses VP2 protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF021235; AAC40994.1; -;
DR

DR InterPro; IPR001742; Orbi_VP2.
DR Pfam; PF00898; Orbi_VP2; 1.
KW Coat protein.
SQ SEQUENCE 1051 AA; 122326 MW; 2B04DB9E389F4B5F CRC64;
Query Match 30.6%; Score 57; DB 1; Length 1051;
Best Local Similarity 47.6%; Pred. No. 4.6;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 1 FNNFTVSFWLRVPKVSASHLE 21
DB 636 FSKRFVSQYRVEKITKHL 656
RESULT 11
ID GON1_RANCA STANDARD; PRT; 90 AA.
AC Q90Y63;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I (GAP1)].
GN GNRH1 OR GNRH.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=6400;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RX TISSUE=Forebrain;
RX MEDLINE=21102951; PubMed=11170016;
RA Wang L., Yoo M.S., Kang H.M., Im W.B., Choi H.S., Bogerd J.,
RA Kwon H.B.;
RT "Cloning and characterization of cDNAs encoding the GnRH1 and GnRH2
RT precursors from bullfrog (Rana catesbeiana).";
RL J. Exp. Zool. 289:190-201(2001).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Forebrain.
CC -!- DEVELOPMENTAL STAGE: Expressed at significantly higher levels
CC during post-breeding. Not expressed in pituitary.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF188754; AAL05972.1; -;
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; NAS.
DR GO; GO:0009755; P:hormone mediated signaling; NAS.
DR GO; GO:0000003; P:reproduction; NAS.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; Gonadoliberin1.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PRO1541; GONADOLIBERN1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Signal;
KW Pyrrolidone carboxylic acid.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 90 PROGNADOLIBERIN I.
FT PEPTIDE 25 34 GONADOLIBERIN I.
FT PEPTIDE 38 86 GNRH-ASSOCIATED PEPTIDE I (BY
FT SIMILARITY).
FT MOD_RES 25 25 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT

```

FT MOD_RES 34 34 AMIDATION (G-35 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
SQ SEQUENCE 90 AA; 10291 MW; 317203B4E3DA2FE7 CRC64;
Query Match 30.4%; Score 56.5; DB 1; Length 90;
Best Local Similarity 46.7%; Pred. No. 0.38;
Matches 14; Conservative 4; Mismatches 9; Indels 3; Gaps 2;
QY 5 TVSFVWLRVKVASH-LEGPSLHWSYGLRP 33
DB 6 TVVLLAIIVLLSSHHHQ- HWSYGLRP 33
RESULT 12
EXE CLOBU STANDARD; PRT; 1250 AA.
AC P30995;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DB Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BoNT/E)
DE (Bontoxilysin E).
OS Clostridium butyricum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1492;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43181, and ATCC 43755;
RX MEDLINE=92181428; PubMed=1543481;
RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RA Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
RA Yokosawa N., Yashiki T., Oguma K.;
RT "Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
RT type E toxin gene from Clostridium butyricum strain BL6340."
RL J. Gen. Microbiol. 137:1519-525(1991).
RN [2]
RP SEQUENCE OF 1-251 FROM N.A.
RC STRAIN=BL6340;
RX MEDLINE=91237316; PubMed=2033376;
RA Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
RA Yokosawa N., Yashiki T., Oguma K.;
RT "Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
RT type E toxin gene from Clostridium butyricum strain BL6340."
RL J. Gen. Microbiol. 137:1519-525(1991).
RN [3]
RP SEQUENCE OF 1-48.
RC STRAIN=5262;
RX Ginepez J., Foley J., Dasgupta B.R.;
RT "Neurotoxin type E from Clostridium botulinum and C. butyricum;
RT partial sequence and comparison."
RL FASEB J. 2:A1750-A1750(1988).
CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC heavy chain (H). The light chain has the pharmacological activity,
CC while the N- and C-terminal of the heavy chain mediate channel
CC formation and toxin binding, respectively.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: there are seven antigenically distinct forms of
CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC -!- SIMILARITY: Belongs to peptidase family M27.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X62088; CAA43998.1; -
CC EMBL; X53180; CAA37321.1; -
CC PIR; JH0256; JH0256.
CC KSP; P10845; 3BTA.
CC MEROPS; M27.002;
CC InterPro; IPR008985; ConA like_lect_gl.
CC InterPro; IPR002160; Kunitz_legume.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR000395; Peptidase_M27.
CC Pfam; PF01742; Peptidase_M27; 1.
CC PRINTS; PR00760; BONTXILYSIN.
CC PRODOM; PD001963; Bontoxilysin; 1.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
KW INIT MET 0 0
FT CHAIN 421 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
FT METAL 211 212 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 212 212 BY SIMILARITY.
FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 411 425 INTERCHAIN (PROBABLE).
FT CONFLICT 229 229 K -> M (IN REF. 2).
SQ SEQUENCE 1250 AA; 143265 MW; 8171B5B2C312857 CRC64;
Query Match 30.4%; Score 56.5; DB 1; Length 1250;
Best Local Similarity 22.4%; Pred. No. 6.5;
Matches 15; Conservative 8; Mismatches 7; Indels 37; Gaps 2;
QY 1 FNNFTVSWLRVVP-----KVSASHLE-----GP 23
DB 911 YKPFISFWRIFNYDNKIVNVEYTIINCMDDNSGKWSLNHEIITWLTQNSGQ 970
QY 24 SLHWSYG 30
DB 971 KLAIFYG 977
RESULT 13
BXA2 CLOBO STANDARD; PRT; 1295 AA.
AC O45854; P77780;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)
DE (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
DE chain; Botulinum neurotoxin A, heavy-chain].
GN BOTA OR BNA OR ATX.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kyoto-F;
RX MEDLINE=94143603; PubMed=8310180;
RX Willems A., East A.K., Lawson P.A., Collins M.D.;
RT "Sequence of the gene coding for the neurotoxin of Clostridium
RT botulinum type A associated with infant botulism: comparison with
RT other clostridial neurotoxins."
RT Res. Microbiol. 144:547-556(1993).
RN [2]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN=Kyoto-F;
RX MEDLINE=97016817; PubMed=8863443;
RX East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding
RT components of the botulinum toxin complex in proteolytic Clostridium

```

RT botulinum types A, B, and F: evidence of chimeric sequences in the
 RL gene encoding the nontoxic nonhemagglutinin component.";
 CC Int. J. Syst. Bacteriol. 46:1105-1112(1996).
 CC -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
 CC binds with high affinity to peripheral neuronal presynaptic
 CC membrane, is then internalized by receptor-mediated endocytosis.
 CC The C-terminus of the heavy chain (H) is responsible for the
 CC adherence of the toxin to the cell surface while the N-terminus
 CC mediates transport of the light chain from the endocytic vesicle
 CC to the cytosol. After translocation, the light chain (L)
 CC hydrolyzes the 197-Gln-198 bond in SNAP-25, thereby blocking
 CC neurotransmitter release. Inhibition of acetylcholine release
 CC results in flaccid paralysis, with frequent heart or respiratory
 CC failure (by similarity).
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
 CC heavy chain (H) (by similarity).
 CC -!- MISCELLANEOUS: There are seven antigenically distinct forms of
 CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
 CC -!- SIMILARITY: Belongs to peptidase family M27.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X73423; CAA51824.1; -;
 CC DR EMBL; X87974; CAA61234.1; -;
 CC DR PIR; I40645; I40645.
 CC DR HSP; P10845; 3BTA.
 CC DR MEROPS; M27.002; -;
 CC DR InterPro; IPR008985; ConsA like lec gl.
 CC DR InterPro; IPR002160; Kunitz legume.
 CC DR InterPro; IPR006025; Pept_M_Zn_BS.
 CC DR InterPro; IPR000395; Peptidase_M27.
 CC DR Pfam; PF01742; Peptidase_M27; 1.
 CC DR PRINTS; PR00760; BONTOXILYSIN.
 CC DR ProDom; PD001963; Bontoxilysin; 1.
 CC DR PROSITE; PS00142; ZINC_PROTEASES; FALSE NEG.
 CC KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 CC INIT MET 0 0 BY SIMILARITY.
 CC FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
 CC FT CHAIN 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
 CC FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
 CC FT ACT_SITE 223 223 BY SIMILARITY.
 CC FT METAL 226 226 ZINC (CATALYTIC) (BY SIMILARITY).
 CC FT DISULFID 429 453 ZINC INTERCHAIN (BY SIMILARITY).
 CC FT DISULFID 1234 1279 BY SIMILARITY.
 CC FT TRANSMEM 626 646 POTENTIAL.
 CC FT TRANSMEM 655 675 POTENTIAL.
 CC SQ SEQUENCE 1295 AA; 149279 MW; 5DA0A13D98D6372 CRC64;
 CC
 CC Query Match 30.1%; Score 56; DB 1; Length 1295;
 CC Best Local Similarity 50.0%; Pred. No. 7.9;
 CC Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 1 ENNFVTSFWLRVPK 14
 CC Db 937 YENFSTFWIKPK 950
 CC
 CC RESULT 14
 CC GONI_HUMAN
 CC ID_GONI_HUMAN STANDARD; PRT; 92 AA.
 CC AC P01146;
 CC DT 21-JUL-1986 (Rel. 01, Created)
 CC DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
 DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 DE hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated
 DE peptide I].
 GN GNRH1 OR GNRH OR LHRH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=89366682; PubMed=2671939;
 RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
 RT "The complete nucleotide sequence of the human gonadotropin-releasing
 RT hormone gene.";
 RL Nucleic Acids Res. 17:6403-6403 (1989).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=86094338; PubMed=2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting
 RT factor in human and rat.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183 (1986).
 RN [3]
 RN SEQUENCE FROM N.A., AND VARIANT SER-16.
 RX MEDLINE=85012739; PubMed=6090951;
 RA Seeburg P.H., Adelman J.P.;
 RT "Characterization of cDNA for precursor of human luteinizing hormone
 RT releasing hormone.";
 RL Nature 311:666-668 (1984).
 RN [4]
 RN SEQUENCE OF 24-33.
 RX MEDLINE=83126573; PubMed=6760865;
 RA Tan L., Rousseau P.;
 RT "The chemical identity of the immunoreactive LHRH-like peptide
 RT biosynthesized in the human placenta.";
 RL Biochem. Biophys. Res. Commun. 109:1061-1071 (1982).
 RN [5]
 RN VARIANT SER-16.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes.";
 RL Nat. Genet. 22:231-238 (1999).
 RN [6]
 RN ERATUM.
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 RA Lander E.S.;
 RL Nat. Genet. 23:373-373 (1999).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
 CC Lutrepulse or Lutrelef (Ferring Pharmaceuticals) and Relisorm
 CC (Serono).
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

```

EMBL; X01059; CAA25525.1; -.
EMBL; M12578; AAA35916.1; -.
EMBL; X15215; CAA33285.1; -.
PIR; S05308; RHUG.
GeneW; HGNC:4419; GNRH1.
MIM; 152760; -.
GO; GO:0005625; C:soluble fraction; TAS.
GO; GO:0005183; F:clustering hormone-releasing factor activity; TAS.
EMBL; X15215; CAA33285.1; -.
PIR; S05308; RHUG.
GeneW; HGNC:4419; GNRH1.
MIM; 152760; -.
GO; GO:0005625; C:soluble fraction; TAS.
GO; GO:0005183; F:clustering hormone-releasing factor activity; TAS.
GO; GO:0007267; P:cell-cell signaling; TAS.
GO; GO:0007275; P:development; TAS.
GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
GO; GO:0007265; P:signal transduction; TAS.
InterPro; IPR002012; GNRH.
InterPro; IPR004079; GonadoliberinI.
Pfam; PF00446; Gnrh; 1.
PRINTS; PS01541; GONADOLIBERIN.
PROSITE; PS00473; GNRH; 1.
Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
Placenta; Pharmaceutical; Signal; Polymorphism;
Pyroglutamate carboxylic acid.
SIGNAL 1 23
CHAIN 24 92
MOD_RES 24 24
MOD_RES 33 33
VARIANT 16 16
SEQUENCE 92 AA; 10380 MW; 30A72221E076FA79 CRC64;

Query Match 29.3%; Score 54.5; DB 1; Length 92;
Best Local Similarity 73.3%; Pred.No. 0.73;
Matches 11; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 20 LEG-PSLHWSYGLRP 33
DB 18 VEGCSSQHSYGLRP 32
::: |||||
::: |||||

RESULT 15
BXCI_CLOBO
ID BXCI_CLOBO STANDARD; PRT; 1290 AA.
AC P18640;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type C1 precursor (EC 3.4.24.69) (BONT/C1)
DE (Bontoxilysin Cl).
DE Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OC NCBI_TaxID=1491;
RN [1]_TaxID=1491;
RN SEQUENCE FROM N.A.
RX MEDLINE=90370487; PubMed=2204031;
RA Hauser D., Eklund M.W., Kurazona H., Binz T., Niemann H., Gill D.M.,
RA Boquet P., Popoff M.R.;
RT "Nucleotide sequence of Clostridium botulinum C1 neurotoxin.";
RN Nucleic Acids Res. 18:4924-4924(1990).
RN [2]
RN SEQUENCE FROM N.A.
RX STRAIN=Type C Stockholm / C-ST;
RX MEDLINE=91024998; PubMed=2222445;
RA Kimura K., Fujii N., Tsuzuki K., Murakami T., Indoh T.,
RA Yokosawa N., Takeshi K., Syuto B., Oguma K.;
RT "The complete nucleotide sequence of the gene coding for botulinum
RT type C1 toxin in the C-ST phage genome.";
RN Biochem. Biophys. Res. Commun. 171:1304-1311(1990).
RN [3]
RN SEQUENCE OF 2-25.
RX STRAIN=Type C Stockholm / C-ST;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 12.0019 seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-9
Perfect score: 160
Sequence: 1 KLLSEIKGVVHRLEGVEGSLHWSYGLRPX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pap:*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pap:*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pap:*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pap:*
5: /cgn2_6/ptodata/2/iaa/PTCUTS-COMB.pap:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 117 | 73.1 | 42 | 1 | US-08-446-692-20 |
| 2 | 117 | 73.1 | 42 | 2 | US-08-488-351A-20 |
| 3 | 116 | 72.5 | 27 | 1 | US-08-446-692-19 |
| 4 | 116 | 72.5 | 27 | 2 | US-08-488-351A-19 |
| 5 | 116 | 72.5 | 27 | 3 | US-09-100-414B-36 |
| 6 | 116 | 72.5 | 27 | 3 | US-09-303-323-36 |
| 7 | 116 | 72.5 | 27 | 4 | US-09-770-014-36 |
| 8 | 116 | 72.5 | 45 | 1 | US-08-446-692-33 |
| 9 | 116 | 72.5 | 45 | 2 | US-08-488-351A-33 |
| 10 | 113 | 70.6 | 27 | 3 | US-09-100-414B-43 |
| 11 | 113 | 70.6 | 27 | 3 | US-09-303-323-43 |
| 12 | 113 | 70.6 | 27 | 4 | US-09-770-014-43 |
| 13 | 110 | 68.8 | 31 | 3 | US-09-100-414B-55 |
| 14 | 110 | 68.8 | 31 | 3 | US-09-303-323-55 |
| 15 | 110 | 68.8 | 31 | 4 | US-09-770-014-55 |
| 16 | 108 | 67.5 | 27 | 3 | US-09-100-414B-41 |
| 17 | 108 | 67.5 | 27 | 3 | US-09-100-414B-47 |
| 18 | 108 | 67.5 | 27 | 3 | US-09-303-323-41 |
| 19 | 108 | 67.5 | 27 | 3 | US-09-303-323-47 |
| 20 | 108 | 67.5 | 27 | 4 | US-08-770-014-41 |
| 21 | 108 | 67.5 | 27 | 4 | US-08-770-014-47 |
| 22 | 108 | 67.5 | 45 | 3 | US-09-100-414B-45 |
| 23 | 108 | 67.5 | 45 | 3 | US-09-303-323-45 |
| 24 | 108 | 67.5 | 45 | 4 | US-09-770-014-45 |
| 25 | 107 | 66.9 | 31 | 3 | US-09-100-414B-59 |
| 26 | 107 | 66.9 | 31 | 3 | US-09-303-323-59 |
| 27 | 107 | 66.9 | 31 | 4 | US-09-770-014-59 |

Sequence 38, Appl
Sequence 38, Appl
Sequence 38, Appl
Sequence 53, Appl
Sequence 53, Appl
Sequence 53, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 57, Appl
Sequence 57, Appl
Sequence 57, Appl
Sequence 17, Appl
Sequence 17, Appl
Sequence 50, Appl
Sequence 50, Appl
Sequence 50, Appl
Sequence 80, Appl

28 106 66.2 28 3 US-09-100-414B-38
29 106 66.2 28 3 US-09-303-323-38
30 106 66.2 28 4 US-09-770-014-38
31 105 65.6 31 3 US-09-100-414B-53
32 105 65.6 31 3 US-09-303-323-53
33 105 65.6 31 4 US-09-770-014-53
34 105 65.6 47 3 US-09-100-414B-60
35 105 65.6 47 3 US-09-303-323-60
36 105 65.6 47 4 US-09-770-014-60
37 105 65.6 49 3 US-09-100-414B-57
38 105 65.6 49 3 US-09-303-323-57
39 105 65.6 49 4 US-09-770-014-57
40 103.5 64.7 25 1 US-08-446-692-17
41 103.5 64.7 25 2 US-08-488-351A-17
42 103 64.4 27 3 US-09-100-414B-50
43 103 64.4 27 3 US-09-303-323-50
44 103 64.4 27 4 US-09-770-014-50
45 103 64.4 35 3 US-09-100-414B-80

ALIGNMENTS

RESULT 1
US-08-446-692-20
; Sequence 20, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-20

Query Match 73.1%; Score 117; DB 1; Length 42;
Best Local Similarity 82.8%; Pred. No. 4.7e-11;
Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 2 LLSIKGVVHRLEGVEGSLHWSYGLRP 30
DB 15 VLSEIKGVVHRLEGVGGE--HWSYGLRP 41

```
RESULT 2
US-08-488-351A-20
; Sequence 20, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-20
Query Match 73.1%; Score 117; DB 2; Length 42;
Best Local Similarity 82.8%; Pred. No. 4.7e-11;
Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 2 LLSEIKGVIVHRLGVGEGSLHWSYGLRP 30
DB 15 VLSEIKGVIVHRLGVGEG--HWSYGLRP 41

RESULT 3
US-08-446-692-19
; Sequence 19, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
```

```
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-19
Query Match 72.5%; Score 116; DB 1; Length 27;
Best Local Similarity 85.7%; Pred. No. 4e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGVGEGSLHWSYGLRP 30
DB 1 LSEIKGVIVHRLGVGEG--HWSYGLRP 26

RESULT 4
US-08-488-351A-19
; Sequence 19, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
```

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/229,275
;; FILING DATE: 14-APR-1994
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/057,166
;; FILING DATE: 27-APR-1992
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Maria C.H. Lin
;; REGISTRATION NUMBER: 29,323
;; REFERENCE/DOCKET NUMBER: 1151-4146 US2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212)415-8745
;; TELEFAX: (516)751-6849
;; INFORMATION FOR SEQ ID NO: 19:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 27 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-488-351A-19

Query Match 72.5%; Score 116; DB 2; Length 27;
Best Local Similarity 85.7%; Pred. No. 4e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30
|||
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 5
US-100-414B-36
; Sequence 36, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-36

Query Match 72.5%; Score 116; DB 3; Length 27;

Best Local Similarity 85.7%; Pred. No. 4e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 3 LSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30
|||
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 6
US-09-303-323-36
; Sequence 36, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-36

Query Match 72.5%; Score 116; DB 3; Length 27;
Best Local Similarity 85.7%; Pred. No. 4e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30
|||
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 7
US-09-770-014-36
; Sequence 36, Application US/09770014
; Patent No. 6559282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York

Query Match 72.5%; Score 116; DB 2; Length 45;
Best Local Similarity 85.7%; Pred. No. 7.3e-11;


```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-770-014-43

```

| Query Match | Score 113; | DB 4; | Length 27; |
|-----------------------|------------|--------------------|--------------|
| Best Local Similarity | 70.6%; | | |
| | 82.1%; | | |
| | | Pred. No. 1.2e-10; | |
| | | | Index |
| | | | 1. Matches |
| | | | 2. Consensus |
| | | | 3. Consensus |

3 LSEIKGVIVHRLEGVGEP SLHWSYGLRP 30
1 LSEIKGVIVHKLEGVGGE --HWSYGLRP 26

RESULT 13
US-09-100-414B-55
; Sequence 55, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:

```

/ GENERAL INFORMATION:
/
/ APPLICANT: Wang, Chang Yi
/
/ TITLE OF INVENTION: NOVEL LHRH PEPTIDE
/
/ TITLE OF INVENTION: IMMUNOGENS
/
/ NUMBER OF SEQUENCES: 106
/

```

ADDRESS: Morgan & Finnegan, L.L.P.

ADDRESS: 102-55
STREET: 345 Park
CITY: New York
STATE: NY
COUNTRY: USA

```

/ COOKING: 500
/
/ ZIP: 10154-0054
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC Windows
/ SOFTWARE: Word 97
/

```

```

/ SOFTWARE:  WOLU 3 /
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER:  US/09/100,414B
/ FILING DATE:  20-JUNE-1998
/ CLASSIFICATION:  424
/ ATTORNEY/AGENT INFORMATION:
/

```

NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```

US-09-100-414B-55

| | | | | |
|-----------------------|-------|------------------|-------|------------|
| Query Match | 68.8% | Score 110; | DB 3; | Length 31; |
| Best Local Similarity | 78.6% | Pred. No. 4e-10; | | |
| Matches | 22; | Conservative | 1; | Mismatches |
| | | | 5; | Indels |

Qy 3 LSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30
 |||||:|||||
 Dh 3 LSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30

RESULT 14
US-09-303-323-55
; Sequence 55, Application US/09303323
; Patent No. 628987
; GENERAL INFORMATION:

```

; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
;

```


ADDRESSEE: Morgan & Finnegan, L.L.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA

ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303.323
FILING DATE: 30-APR-1999

FILING DATE: 20 JUN 1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA: 09/100,414
 APPLICATION NUMBER: 09/100,414
 FILING DATE: 20-JUNE-1998
 ATTORNEY/AGENT INFORMATION:

NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800

TELEPHONE: 212-751-6849
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
-303-323-55

ry Match 68.8%; Score 110; DB 3; Length 31;
 t Local Similarity 78.6%; Pred. No. 4e-10;
 ches 22; Conservative 1; Mismatches 5; Indels

3 LSEIKGVIVHRLEGVGPSLHWSYGLRP 30
3 LSEIKGVIVHKLEGLFGGEHWSYGLRP 30

```

RESULT 15
US-09-7700-014-55
; Sequence 55, Application US/097700:4
; Patent No. 6559282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
;

```

ADDRESS: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323

```

; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-770-014-55

```

```

Query Match      68.8%; Score 110; DB 4; Length 31;
Best Local Similarity 78.6%; Pred. No. 4e-10;
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      3 LSEIKGVIVHRLGVEGFSLHWSYGLRP 30
DB      3 LSEIKGVIVHRLGVEGFSLHWSYGLRP 30

```

```

Search completed: March 10, 2004, 09:28:53
Job time : 12.0019 secs

```



```

; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
; FILE REFERENCE: Immunogen Delivery System
; CURRENT APPLICATION NUMBER: US/10/076,674
; CURRENT FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Human
; US-10-076-674-9
;
Query Match 72.5%; Score 116; DB 14; Length 45;
Best Local Similarity 85.7%; Pred. No. 1.5e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
;
QY 3 LSEIKGVIVHRLGVGSPSLHWSYGLRP 30
      |||||
DB 19 LSEIKGVIVHRLGVGGE--HWSYGLRP 44
      |||||
;
RESULT 4
; US-10-355-161A-9
; Sequence 9, Application US/10355161A
; Publication No. US20040009897A1
; GENERAL INFORMATION:
; APPLICANT: Sokoll, Kenneth K.
; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
; FILE REFERENCE: Immunogen Delivery System
; CURRENT APPLICATION NUMBER: US/10/355,161A
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 10/076674
; PRIOR FILING DATE: 2002-02-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Human
; US-10-355-161A-9
;
Query Match 72.5%; Score 116; DB 15; Length 45;
Best Local Similarity 85.7%; Pred. No. 1.5e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
;
QY 3 LSEIKGVIVHRLGVGSPSLHWSYGLRP 30
      |||||
DB 19 LSEIKGVIVHRLGVGGE--HWSYGLRP 44
      |||||
;
RESULT 5
; US-09-847-102A-33
; Sequence 33, Application US/09847102A
; Publication No. US2003004409A1
; GENERAL INFORMATION:
; APPLICANT: University of California
; APPLICANT: Carson, Dennis A.
; APPLICANT: Corr, Maripat
; APPLICANT: Rhee, Chae-Seo
; APPLICANT: Lorenzo, Leoni M.
; APPLICANT: Malini, Sen
; TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS
; CURRENT APPLICATION NUMBER: US/09/847,102A
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

```

```
; OTHER INFORMATION: PMVFP-FZD2
US-09-847-102A-33
Query Match 61.9%; Score 99; DB 10; Length 75;
Best Local Similarity 95.5%; Pred. No. 7.7e-07;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLSEIKGVVHRLEGVEGPSL 22
    |||||
Db 2 KLLSLIKGVVHRLEGVEGPSL 23
    |||||

RESULT 6
US-10-285-976-231
; Sequence 231, Application US/10285976
; Publication No. US20030165500A1
; GENERAL INFORMATION:
; APPLICANT: Rhes, Chae-Seo
; APPLICANT: Malini, Sen
; APPLICANT: Wu, Christina
; APPLICANT: Leoni, Lorenzo M.
; APPLICANT: Carr, Maripat
; APPLICANT: Carson, Dennis A.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy
; TITLE OF INVENTION: In Head and Neck Squamous Cell Carcinomas
; FILE REFERENCE: 023070-130320US
; CURRENT APPLICATION NUMBER: US/10/285,976
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US 60/287,995
; PRIOR FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: WO PCT/US02/13802
; PRIOR FILING DATE: 2002-05-01
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 231
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PMVFP-ZD2
; OTHER INFORMATION: measles virus fusion (MVF) epitope fused to
; OTHER INFORMATION: frizzled domain
US-10-285-976-231

Query Match 61.9%; Score 99; DB 14; Length 75;
Best Local Similarity 95.5%; Pred. No. 7.7e-07;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLSEIKGVVHRLEGVEGPSL 22
    |||||
Db 2 KLLSLIKGVVHRLEGVEGPSL 23
    |||||

RESULT 7
US-09-848-834A-13
; Sequence 13, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 13
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 947-967 of the Tetanus Toxoid Precursor
; OTHER INFORMATION: Tetanus toxoid precursor (Tetoxylysin) linked by a spacer to a
; OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; NAME/KEY: PEPTIDE
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(21)
; OTHER INFORMATION: Amidated phenylalanine
; LOCATION: (1)..(1)
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(21)
; OTHER INFORMATION: Amino acids 947-967 of the Tetanus Toxoid Precursor
; OTHER INFORMATION: (Tetoxylysin)
; NAME/KEY: PEPTIDE
; LOCATION: (22)..(25)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (26)..(34)
; OTHER INFORMATION: Amino acids 2-10 of the human GnRH hormone
; NAME/KEY: MOD RES
; LOCATION: (34)..(34)
; OTHER INFORMATION: Amidated glycine or glycineamide
US-09-848-834A-10

Query Match 49.4%; Score 79; DB 9; Length 34;
Best Local Similarity 92.9%; Pred. No. 0.00025;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 17 VEGPSLHWSYGLRP 30
    :|||||
```

Db 20 LEGPSLHWSYGLRP 33

RESULT 9

US-09-848-834A-18

Sequence 18, Application US/09848834A

Patent No. US20020076416A1

GENERAL INFORMATION:

APPLICANT: Aptton Corporation

TITLE OF INVENTION: Chimeric Peptide Immunogens

FILE REFERENCE: 1102865-0047

CURRENT APPLICATION NUMBER: US/09/848,834A

CURRENT FILING DATE: 2001-05-04

PRIOR APPLICATION NUMBER: 60/202,328

PRIOR FILING DATE: 2000-05-05

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patentin version 3.0

SEQ ID NO 18

LENGTH: 50

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human GnRH linked by a spacer to amino acid sequence 947-967 of the Tetanus toxin precursor (Tentoxigenin) protein linked by a spacer to amino acid sequence 2-10 of human GnRH

NAME/KEY: MOD RES

LOCATION: (1)..(1)

OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

NAME/KEY: MOD RES

LOCATION: (50)..(50)

OTHER INFORMATION: Amidated glycine or glycineamide

NAME/KEY: PEPTIDE

LOCATION: (1)..(10)

OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone

NAME/KEY: PEPTIDE

LOCATION: (11)..(16)

OTHER INFORMATION: Spacer peptide

NAME/KEY: PEPTIDE

LOCATION: (17)..(37)

OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxin precursor (Tentoxigenin)

NAME/KEY: PEPTIDE

LOCATION: (38)..(41)

OTHER INFORMATION: Spacer peptide

NAME/KEY: PEPTIDE

LOCATION: (42)..(50)

OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone

US-09-848-834A-18

Query Match 49.4%; Score 79; DB 9; Length 50;

Best Local Similarity 52.9%; Pred. No. 0.00039;

Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 17 VEGPSLHWSYGLRP 30

Db 36 LEGPSLHWSYGLRP 49

RESULT 10

US-10-223-711-10

Sequence 10, Application US/10223711

Publication No. US20030113344A1

GENERAL INFORMATION:

APPLICANT: Bakaletz, Lauren O.

TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides

FILE REFERENCE: 18525/04058

CURRENT APPLICATION NUMBER: US/10/223,711

CURRENT FILING DATE: 2002-08-19

PRIOR APPLICATION NUMBER: 09/148,711

PRIOR FILING DATE: 1998-09-04

PRIOR APPLICATION NUMBER: 08/460,502

PRIOR FILING DATE: 1995-06-02

NUMBER OF SEQ ID NOS: 12

SOFTWARE: Patentin version 3.1

SEQ ID NO 10

LENGTH: 40

TYPE: PRT

ORGANISM: Artificial

FEATURE:

OTHER INFORMATION: synthetic construct

US-10-223-711-10

Query Match 48.8%; Score 78; DB 14; Length 40;

Best Local Similarity 94.4%; Pred. No. 0.00042;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLSEIKGVIVHRLGVE 18

Db 23 KLLSLIKGVIVHRLGVE 40

RESULT 11

US-09-847-102A-31

Sequence 31, Application US/09847102A

Publication No. US2003004409A1

GENERAL INFORMATION:

APPLICANT: University of California

APPLICANT: Carson, Dennis A.

APPLICANT: Carr, Maripat

APPLICANT: Rhee, Chae-Seo

APPLICANT: Lorenzo, Leoni M.

APPLICANT: Malini, Sen

TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS

TITLE OF INVENTION: STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS

FILE REFERENCE: 22000-20629.00

CURRENT APPLICATION NUMBER: US/09/847,102A

CURRENT FILING DATE: 2001-05-01

NUMBER OF SEQ ID NOS: 138

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 31

LENGTH: 75

TYPE: PRT

ORGANISM: Artificial sequence

FEATURE:

OTHER INFORMATION: PFZD2-MMVF

US-09-847-102A-31

Query Match 48.8%; Score 78; DB 10; Length 75;

Best Local Similarity 94.4%; Pred. No. 0.00085;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLSEIKGVIVHRLGVE 18

Db 58 KLLSLIKGVIVHRLGVE 75

RESULT 12

US-10-285-976-229

Sequence 229, Application US/10285976

Publication No. US20030165500A1

GENERAL INFORMATION:

APPLICANT: Rhee, Chae-Seo

APPLICANT: Malini, Sen

APPLICANT: Wu, Christina

APPLICANT: Leoni, Lorenzo M.

APPLICANT: Carr, Maripat

APPLICANT: Carson, Dennis A.

APPLICANT: The Regents of the University of California

TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy

TITLE OF INVENTION: In Head and Neck Squamous Cell Carcinomas

FILE REFERENCE: 023070-130320US

CURRENT APPLICATION NUMBER: US/10/285,976

CURRENT FILING DATE: 2002-11-01

PRIOR APPLICATION NUMBER: 60/287,995

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 9.64981 Seconds
(without alignments)
309.015 Million cell updates/sec

Title: US-09-848-834A-9
Perfect score: 100
Sequence: 1 KLSSEIKGVIVHRLGVVSLHWSYGLRPX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 72 | 45.0 | 282 | 2 P00376 | cell fusion glycop |
| 2 | 72 | 45.0 | 282 | 2 P00388 | cell fusion glycop |
| 3 | 72 | 45.0 | 534 | 1 JU0274 | cell fusion glycop |
| 4 | 72 | 45.0 | 546 | 2 S47300 | gene F protein - r |
| 5 | 72 | 45.0 | 550 | 1 E48556 | cell fusion glycop |
| 6 | 72 | 45.0 | 553 | 1 VGNZMW | cell fusion glycop |
| 7 | 71 | 44.4 | 546 | 1 VGNZRX | cell fusion glycop |
| 8 | 71 | 44.4 | 546 | 2 S47305 | gene F protein - r |
| 9 | 66.5 | 41.6 | 552 | 2 S47034 | cell fusion glycop |
| 10 | 66 | 41.2 | 546 | 1 VGNZRL | cell fusion glycop |
| 11 | 65 | 40.6 | 542 | 2 J02223 | cell fusion glycop |
| 12 | 65 | 40.6 | 662 | 1 VGNZCD | cell fusion glycop |
| 13 | 64 | 40.6 | 682 | 2 S21382 | cell fusion glycop |
| 14 | 64 | 40.0 | 631 | 1 VGNZPD | cell fusion glycop |
| 15 | 64 | 40.0 | 631 | 1 A48346 | cell fusion glycop |
| 16 | 60 | 37.5 | 546 | 2 S55386 | cell fusion glycop |
| 17 | 57.5 | 35.9 | 92 | 1 RHUG | gonadolibirin prec |
| 18 | 57 | 35.6 | 90 | 1 RHUG | gonadolibirin prec |
| 19 | 57 | 35.6 | 92 | 1 RHUG | gonadolibirin prec |
| 20 | 56 | 35.0 | 636 | 2 S47299 | gene F protein - r |
| 21 | 54 | 33.8 | 67 | 2 I78541 | gonadolibirin prec |
| 22 | 53 | 33.1 | 508 | 1 VGNVFR | spike glycoprotein |
| 23 | 52 | 32.5 | 10 | 1 RHPG | gonadolibirin - pi |
| 24 | 52 | 32.5 | 10 | 1 RHPG | gonadolibirin - sh |
| 25 | 52 | 32.5 | 89 | 2 I51423 | gonadolibirin prec |
| 26 | 52 | 32.5 | 379 | 1 D8SPX | phosphoribosylamin |
| 27 | 50.5 | 31.6 | 190 | 2 T37168 | probable tetr-fam1 |
| 28 | 50 | 31.2 | 333 | 2 T3151 | hypothetical prote |
| 29 | 49 | 30.6 | 451 | 2 AH0063 | conserved hypothet |

30 49 30.6 582 2 G71500
31 48.5 30.3 393 2 D83589
32 48.5 30.3 413 2 B70907
33 48 30.0 10 1 RHAQ1
34 48 30.0 92 2 I50644
35 48 30.0 98 2 I50739
36 48 30.0 502 2 T36589
37 48 30.0 576 2 AF2361
38 48 30.0 856 2 I58411
39 47.5 29.7 222 1 E71024
40 47.5 29.7 229 2 JG7219
41 47.5 29.7 586 2 T08293
42 47 29.4 339 2 B97755
43 47 29.4 388 2 C72710
44 47 29.4 452 2 T10882
45 47 29.4 516 2 T00514

aspartate-tRNA lig
glutaryl-CoA dehyd
hypothetical prote
gonadolibirin I -
gonadolibirin I pr
gonadotropin-relea
probable transmem
flavoprotein [impo
protein-tyrosine k
probable transport
nuclear protein SR
hypothetical prote
hypothetical prote
probable fmu prote
coproporphyrinogen
cytochrome P450 ho

ALIGNMENTS

RESULT 1

P00376
cell fusion glycoprotein - measles virus (strain TT) (fragment)
C:Species: measles virus
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999
C:Accession: P00376
R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A:Title: A measles virus isolate from a child with Kawasaki disease: sequence comparis
A:Reference number: P00374; MUID:92300360; PMID:1607874
A:Accession: P00376
A:Molecule type: genomic RNA
A:Residues: 1-282 <SCH>
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion

Query Match 45.0%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGV 17

Db 20 LSEIKGVIVHRLGV 34

RESULT 2

P00388
cell fusion glycoprotein - measles virus (strain Schwarz vaccine) (fragment)
C:Species: measles virus
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999
C:Accession: P00388
R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A:Title: A measles virus isolate from a child with Kawasaki disease: sequence comparis
A:Reference number: P00374; MUID:92300360; PMID:1607874
A:Accession: P00388
A:Molecule type: genomic RNA
A:Residues: 1-282 <SCH>
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion

Query Match 45.0%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGV 17

Db 20 LSEIKGVIVHRLGV 34

```

RESULT 3
JU0274
cell fusion glycoprotein precursor - subacute sclerosing panencephalitis virus (strain Y
N1)
C:Species: fusion glycoprotein F1; fusion glycoprotein F2
C:Species: subacute sclerosing panencephalitis virus, SSPV
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jun-2000
C:Accession: JU0274
R:Komase, K.; Haga, T.; Yoshikawa, Y.; Sato, T.A.; Yamanouchi, K.
Virus Genes 4, 173-181, 1990
A:Title: Molecular analysis of structural protein genes of the Yamagata-1 strain of defe
A:Reference number: JU0274; MUID:90385702; PMID:1698327
A:Accession: JU0274
A:Molecule type: mRNA
A:Residues: 1-534 <COM>
A:Cross-references: EMBL:DI0548; NID:9222257; PID:G222257
A:Note: the authors translated the codon GTA for residue 459 as Gly and GGG for residue
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-107/Product: cell fusion glycoprotein F2 #status predicted <PF2>
F:108-534/Product: cell fusion glycoprotein F1 #status predicted <PF1>
F:498-514/Domain: transmembrane #status predicted <TMN>
F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 45.0%; Score 72; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVVHRLEGV 17
DB 288 LSEIKGVVHRLEGV 302

RESULT 4
S47300
gene F protein - rinderpest virus
C:Species: rinderpest virus
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
C:Accession: S47300; PQ0865
R:Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.
submitted to the EMBL Data Library, March 1994
A:Description: The complete nucleotide sequence of the fusion protein gene of the vaccin
A:Reference number: S47299
A:Accession: S47300
A:Molecule type: DNA
A:Residues: 1-546 <EVA>
A:Cross-references: EMBL:Z31656; NID:9535406; PIDN:CAA83482.1; PID:G535407
R:Chamberlain, R.W.; Wanwayi, H.M.; Hockley, E.; Shaila, M.S.; Goatley, L.; Knowles, N.J.
J. Gen. Virol. 74, 2775-2780, 1993
A:Title: Evidence for different lineages of rinderpest virus reflecting their geographic
A:Reference number: PQ0865; MUID:94103786; PMID:8277286
A:Accession: PQ0865
A:Molecule type: mRNA
A:Residues: 86-191 <CHA>
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
Query Match 45.0%; Score 72; DB 2; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVVHRLEGV 17
DB 284 LSEIKGVVHRLEGV 298

RESULT 5

```

```

E48556
cell fusion glycoprotein precursor - measles virus (strain AIK-C)
C:Species: measles virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
C:Accession: E48556
R:Mori, T.; Sasaki, K.; Hashimoto, H.; Makino, S.
Virus Genes 7, 67-81, 1993
A:Title: Molecular cloning and complete nucleotide sequence of genomic RNA of the AIK-C
A:Reference number: A48556; MUID:93227570; PMID:8470368
A:Accession: E48556
A:Molecule type: genomic RNA
A:Residues: 1-550 <WOR>
A:Cross-references: GB:S58435; NID:9299460; PIDN:AA826145.1; PID:G299465
A:Note: sequence extracted from NCBI backbone (NCBI:129264; NCBIP:129272)
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-107/Product: cell fusion glycoprotein F2 #status predicted <PF2>
F:108-550/Product: cell fusion glycoprotein F1 #status predicted <PF1>
F:113-138/Region: hydrophobic
F:495-514/Domain: transmembrane #status predicted <TMN>
F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 45.0%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVVHRLEGV 17
DB 288 LSEIKGVVHRLEGV 302

RESULT 6
VGNZMV
cell fusion glycoprotein precursor - measles virus
C:Species: measles virus
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1989 #text_change 16-Jun-2000
C:Accession: A26962; A25616; PQ0384
R:Buckland, R.; Gerald, C.; Barker, R.; Wild, T.F.
J. Gen. Virol. 68, 1695-1703, 1987
A:Title: Fusion glycoprotein of measles virus: nucleotide sequence of the gene and com
A:Reference number: A26962; MUID:87224816; PMID:3585281
A:Accession: A26962
A:Molecule type: mRNA
A:Residues: 1-553 <BUG>
A:Cross-references: GB:D00090; NID:G222061; PIDN:BAA00056.1; PID:G222062
R:Richardson, C.; Hull, D.; Greer, P.; Hasel, K.; Berkovich, A.; Englund, G.; Bellini,
Virology 155, 508-523, 1986
A:Title: The nucleotide sequence of the mRNA encoding the fusion protein of measles vir
A:Reference number: A94350; MUID:87071668; PMID:3788062
A:Accession: A25616
A:Molecule type: mRNA
A:Residues: 4-553 <RIC>
A:Cross-references: GB:M14915; NID:G331762; PIDN:AAA46423.1; PID:G331763
R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A:Title: A measles virus isolate from a child with Kawasaki disease: sequence comparis
A:Reference number: PQ0374; MUID:92300360; PMID:1607874
A:Accession: PQ0380
A:Molecule type: genomic RNA
A:Residues: 272-553 <SCH1>
A:Experimental source: isolate CL
A:Accession: PQ0384
A:Molecule type: genomic RNA
A:Residues: 272-553 <SCH2>
A:Experimental source: isolate SE
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein

```

C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-110/Product: cell fusion glycoprotein F2 #status predicted <PF2>
F;111-553/Product: cell fusion glycoprotein F1 #status predicted <PF1>
F;501-517/Domain: transmembrane #status predicted <TMN>
F;32,64,70/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 45.0%; Score 72; DB 1; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
|||||:|||||

Db 291 LSEIKGVIVHRLEGV 305
|||||:|||||

RESULT 7
VGNZK
cell fusion glycoprotein precursor - rinderpest virus (strain Kabete O)
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2
C;Species: rinderpest virus
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 25-Oct-1996
C;Accession: A31051
R;Hsu, D.; Yamanaka, M.; Miller, J.; Dale, B.; Grubman, M.; Yilma, T.
Virology 166, 149-153, 1988
A;Title: Cloning of the fusion gene of rinderpest virus: comparative sequence analysis
A;Reference number: A31051; MUID:88322864; PMID:3413983
A;Accession: A31051
A;Molecule type: genomic RNA
A;Residues: 1-546 <HSU>
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-108/Product: cell fusion glycoprotein F2 #status predicted <PF1>
F;109-546/Product: cell fusion glycoprotein F1 #status predicted <PF2>
F;109-134/Domain: transmembrane #status predicted <TMN>
F;491-513/Domain: transmembrane #status predicted <TN2>
F;25,57,63,518/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 44.4%; Score 71; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.046;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
|||||:|||||

Db 284 LSEIKGVIVHRLEGV 298
|||||:|||||

RESULT 8
S47305
gene F protein - rinderpest virus
C;Species: rinderpest virus
C;Date: 20-Oct-1984 #sequence_revision 08-Sep-1995 #text_change 20-Sep-1999
C;Accession: S47305; S47301
R;Baron, M.D.; Barrett, T.
submitted to the EMBL Data Library, March 1994
A;Description: The sequence of the N and L genes of Rinderpest virus, and the 50 and 30
A;Reference number: S47283
A;Accession: S47305
A;Molecule type: mRNA
A;Residues: 1-546 <BAR>
A;Cross-references: EMBL:Z30697; NID:G535396; PIDN:CAA83181.1; PID:G535401; EMBL:Z30700;
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: transmembrane protein

Query Match 44.4%; Score 71; DB 2; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.046;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
|||||:|||||

Db 284 LSEIKGVIVHRLEGV 298
|||||:|||||

RESULT 9
S47034
cell fusion protein precursor - porpoise morbillivirus
N;Alternate names: F protein
C;Species: porpoise morbillivirus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C;Accession: S47034
R;Bolt, G.; Gottschalk, E.; Blixenkron-Moeller, M.; Wisaupt, R.G.A.; Welsh, M.J.;
submitted to the EMBL Data Library, July 1994
A;Description: Nucleotide sequence comparisons of the F and M genes of cetacean morbill
A;Reference number: S47034
A;Accession: S47034
A;Molecule type: mRNA
A;Residues: 1-552 <BOL>
A;Cross-references: EMBL:X80757; NID:G520639; PIDN:CAA56731.1; PID:G520640
A;Experimental source: isolate Ulster 88
A;Note: the source is designated as Cetacean morbillivirus
C;Superfamily: parainfluenza virus cell fusion protein
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-552/Product: fusion protein #status predicted <MAP>

Query Match 41.6%; Score 66.5; DB 2; Length 552;
Best Local Similarity 61.5%; Pred. No. 0.2;
Matches 16; Conservative 1; Mismatches 2; Indels 7; Gaps 1;

QY 3 LSEIKGVIVHRLEGVGPSLHWSYGL 28
|||||:|||||

Db 290 LSEIKGVIVHRLEAV-----SYNL 308
|||||:|||||

RESULT 10
VGNZRL
cell fusion glycoprotein precursor - rinderpest virus (strain L)
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2
C;Species: rinderpest virus
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jul-1999
C;Accession: A28921
R;Tsukiyama, K.; Yoshikawa, Y.; Yamanouchi, K.
Virology 164, 523-530, 1988
A;Title: Fusion glycoprotein (F) of rinderpest virus: entire nucleotide sequence of th
A;Reference number: A28921; MUID:88219541; PMID:3285575
A;Accession: A28921
A;Molecule type: mRNA
A;Residues: 1-546 <TSU>
A;Cross-references: GB:M20870; NID:G333898; PIDN:AAA47399.1; PID:G333899
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-104/Product: cell fusion glycoprotein F2 #status predicted <FG2>
F;105-546/Product: cell fusion glycoprotein F1 #status predicted <FG1>
F;109-133/Domain: transmembrane #status predicted <TMN>
F;485-513/Domain: transmembrane #status predicted <TN2>
F;25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.2%; Score 66; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.23;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
|||||:|||||

Db 284 LSEIKGVIVHRLESV 298
|||||:|||||

RESULT 11
JQ2223
cell fusion protein F0 precursor - phocine distemper virus
N;Contains: F1 and F2 chains
C;Species: phocine distemper virus

C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 24-Nov-1999
C/Accession: JQ2223
R/Visser, I.K.G.; van der Heijden, R.W.J.; van de Bildt, M.W.G.; Kenter, M.J.H.; Oervell
J. Gen. Virol. 74, 1989-1994, 1993
A/Title: Fusion protein gene nucleotide sequence similarities, shared antigenic sites an
e virus entity.
A/Reference number: JQ2223; MUID:93389459; PMID:8376973
A/Accession: JQ2223
A/Molecule type: mRNA
A/Residues: 1542 <VIS>
A/Cross-references: GB107075
A/Note: the authors translated the codon ATC for residue 4 as Leu
C/Comment: This fusion protein F0 is cleaved into F1 and F2 chains.
C/Genetics:
A/Gene: F
C/Superfamily: parainfluenza virus cell fusion protein
C/Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-542/Product: fusion protein #status predicted <MAT>
F:16-99/Product: F2 chain #status predicted <F2C>
F:105-542/Product: F1 chain #status predicted <F1C>
F:105-135/Region: hydrophobic
F:486-512/Domain: transmembrane #status predicted <TMW>
F:21,53,59,397/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 40.6%; Score 65; DB 2; Length 542;
Best Local Similarity 86.7%; Pred. No. 0.32; Mismatches 1; Indels 0; Gaps 0;
Matches 13; Conservative 1;
QY 3 LSEIKGVIVHRLGV 17
|||:|||||
Db 280 LSEVKGIVVHRLVAV 294
|||:|||||
RESULT 12
VGNZCD
cell fusion glycoprotein precursor - canine distemper virus
N/Contains: fusion protein F1; fusion protein F2
C/Species: canine distemper virus
C/Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C/Accession: JS0321
R/Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.
Virus Res. 8, 373-386, 1987
A/Title: The nucleotide sequence of the gene encoding the F protein of canine distemper
A/Reference number: JS0321; MUID:88129050; PMID:3433924
A/Accession: JS0321
A/Molecule type: mRNA
A/Residues: 1-662 <BAR>
A/Cross-references: GB:M21849; NID:G323241; PIDN:AAA42878.1; PID:G323242
C/Genetics:
A/Gene: F
C/Superfamily: parainfluenza virus cell fusion protein
C/Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-135/Domain: signal sequence #status predicted <SIG>
F:136-224/Product: call fusion glycoprotein F2 #status predicted <F2P>
F:225-662/Product: call fusion glycoprotein F1 #status predicted <F1P>
F:606-629/Domain: transmembrane #status predicted <MEM>
F:62,141,173,179,517/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 40.6%; Score 65; DB 1; Length 662;
Best Local Similarity 86.7%; Pred. No. 0.4; Mismatches 1; Indels 0; Gaps 0;
Matches 13; Conservative 1;
QY 3 LSEIKGVIVHRLGV 17
|||:|||||
Db 400 LSEVKGIVVHRLVAV 414
|||:|||||
RESULT 13
S21382
cell fusion protein - canine distemper virus
C/Species: canine distemper virus
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999

C/Accession: S21382
R/Wild, T.F.; Bernard, A.; Spöhner, D.; Villevall, D.; Drillien, R.
submitted to the EMBL Data Library, April 1992
A/Description: Vaccination of mice against canine distemper virus induced encephalitis
A/Reference number: S21382
A/Accession: S21382
A/Status: preliminary
A/Molecule type: genomic RNA
A/Residues: 1-662 <WIL>
A/Cross-references: EMBL:X65509; NID:G58853; PIDN:CAA46481.1; PID:G58854
C/Superfamily: parainfluenza virus cell fusion protein
Query Match 40.6%; Score 65; DB 2; Length 662;
Best Local Similarity 86.7%; Pred. No. 0.4; Mismatches 1; Indels 0; Gaps 0;
Matches 13; Conservative 1;
QY 3 LSEIKGVIVHRLGV 17
|||:|||||
Db 400 LSEVKGIVVHRLVAV 414
|||:|||||
RESULT 14
VGNZPD
cell fusion glycoprotein precursor - phocine distemper virus
N/Contains: fusion protein F1; fusion protein F2
C/Species: phocine distemper virus
C/Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 25-Oct-1996
C/Accession: JQ1368
R/Koenevaes, J.; Blixenkron-Moeller, M.; Sharma, B.; Oervell, C.; Norby, E.
J. Gen. Virol. 72, 2959-2966, 1991
A/Title: The nucleotide sequence and deduced amino acid composition of the haemagglutin
A/Reference number: JQ1368; MUID:92113538; PMID:1765768
A/Accession: JQ1368
A/Molecule type: genomic RNA
A/Residues: 1-631 <KOV>
C/Genetics:
A/Gene: F
C/Superfamily: parainfluenza virus cell fusion protein
C/Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
F:89-106/Domain: transmembrane #status predicted <TM1>
F:189-193/Region: cleavage processing #status predicted
F:194-631/Product: call fusion glycoprotein F1 #status predicted <FP1>
F:194-212/Domain: transmembrane #status predicted <TM2>
F:575-595/Domain: transmembrane #status predicted <TM3>
F:110,142,148,486/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 40.0%; Score 64; DB 1; Length 631;
Best Local Similarity 80.0%; Pred. No. 0.52; Mismatches 2; Indels 0; Gaps 0;
Matches 12; Conservative 2;
QY 3 LSEIKGVIVHRLGV 17
|||:|||||
Db 369 LSEVKGIVVHRLVAV 383
|||:|||||
RESULT 15
A48346
cell fusion glycoprotein precursor - phocine distemper virus (strain Ulster/88)
N/Contains: fusion protein F1; fusion protein F2
C/Species: phocine distemper virus
C/Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 25-Oct-1996
C/Accession: A48346
R/Curran, M.D.; Lu, Y.J.; Rima, B.K.
Arch. Virol. 126, 159-169, 1992
A/Title: The fusion protein gene of phocine distemper virus: nucleotide and deduced ami
A/Reference number: A48346; MUID:92398437; PMID:1524494
A/Accession: A48346
A/Molecule type: mRNA
A/Residues: 1-631 <CUR>
A/Note: sequence extracted from NCBI backbone (NCBIN:113098, NCBI:P:113099)
C/Genetics:
A/Gene: F

C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
F:89-106/Domain: transmembrane #status predicted <TM1>
F:194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>
F:194-219/Domain: transmembrane #status predicted <TM2>
F:575-595/Domain: transmembrane #status predicted <TM3>
F:110,142/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.0%; Score 64; DB 1; Length 631;
Best Local Similarity 80.0%; Pred. No. 0.52;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LSEIKGVVHRLEGV 17
|||:|||||
Db 369 LSEVKGVVVHRLAV 383

Search completed: March 10, 2004, 09:16:45
Job time : 10.7086 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 5.66926 Seconds
(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-9

Perfect score: 160
Sequence: 1 KLLSEIKGVIVHRLGEGPSLHWSYGLRFX 31

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 72 | 45.0 | 534 | 1 VGLF_MEASY | P26032 measles vir |
| 2 | 72 | 45.0 | 546 | 1 VGLF_RINDP | P41360 rinderpest |
| 3 | 72 | 45.0 | 550 | 1 VGLF_MEASA | P35973 measles vir |
| 4 | 72 | 45.0 | 550 | 1 VGLF_MEASE | P08300 measles vir |
| 5 | 71 | 44.4 | 546 | 1 VGLF_RINDR | P41356 rinderpest |
| 6 | 66 | 41.2 | 546 | 1 VGLF_RINDL | P10864 rinderpest |
| 7 | 65 | 40.6 | 546 | 1 VGLF_RINDK | P12574 rinderpest |
| 8 | 65 | 40.6 | 662 | 1 VGLF_CDVO | P12589 canine dist |
| 9 | 64 | 40.0 | 529 | 1 VGLF_MEASI | P26031 measles vir |
| 10 | 64 | 40.0 | 631 | 1 VGLF_PHODV | P28886 phocine dis |
| 11 | 57.5 | 35.9 | 91 | 1 GONI_PIG | P49921 sus scrofa |
| 12 | 57.5 | 35.9 | 92 | 1 GONI_HUMAN | P01148 homo sapien |
| 13 | 57 | 35.6 | 90 | 1 GONI_MOUSE | P13562 mus musculu |
| 14 | 57 | 35.6 | 92 | 1 GONI_RAT | P07490 rattus norv |
| 15 | 54 | 33.8 | 67 | 1 GONI_MACMU | P55247 macaca mula |
| 16 | 53 | 33.1 | 508 | 1 VGLF_THNV | P07923 infectious |
| 17 | 52.5 | 32.8 | 90 | 1 GONI_RANCA | P09063 rana catesb |
| 18 | 52 | 32.5 | 61 | 1 GONI_SHEEP | P02858 ovis aries |
| 19 | 52 | 32.5 | 63 | 1 GONI_MESAU | P09163 mesocricetu |
| 20 | 52 | 32.5 | 89 | 1 GONI_XENLA | P45656 xenopus lae |
| 21 | 52 | 32.5 | 92 | 1 GONI_TURGS | P05335 tupia glis |
| 22 | 52 | 32.5 | 379 | 1 PURK_BACSU | P12045 bacillus su |
| 23 | 50.5 | 31.6 | 92 | 1 GONI_CAVPO | P04713 cavia porce |
| 24 | 50.5 | 31.6 | 99 | 1 GONI_DICLA | P09140 dicentrarch |
| 25 | 49 | 30.6 | 169 | 1 CX41_THUOB | P04840 thunnus obe |
| 26 | 48 | 30.6 | 582 | 1 SYD_CHLTR | P04546 chlamydia t |
| 27 | 48.5 | 30.3 | 95 | 1 GONI_MORSA | P07812 morone saxa |
| 28 | 48 | 30.0 | 10 | 1 GONI_ALLMI | P37041 alligator m |
| 29 | 48 | 30.0 | 92 | 1 GONI_CHICK | P37042 gallus gall |
| 30 | 48 | 30.0 | 94 | 1 GONI_HAPBU | P19184 haplochromi |
| 31 | 48 | 30.0 | 95 | 1 GONI_PAGNA | P70074 pagrus majo |
| 32 | 48 | 30.0 | 95 | 1 GONI_SPAU | P51919 sparus auri |
| 33 | 48 | 30.0 | 124 | 1 UCN1_HUMAN | P55089 homo sapien |

34 48 30.0 576 1 DPA1_ANASP
35 47 29.4 122 1 UCN1_MOUSE
36 47 29.4 452 1 HEMP_RHOSH
37 47 29.4 452 1 HEMP_RHOSH
38 47 29.4 573 1 DPA1_SYNX3
39 47 29.4 579 1 DPA1_ANASP
40 46.5 29.1 207 1 TER2_ECOLI
41 46.5 29.1 417 1 RFI_PYRAB
42 46 28.7 110 1 YHB1_ACTAC
43 46 28.7 555 1 VGLF_P1LHC
44 46 28.7 652 1 GSH1_ONCVO
45 45.5 28.4 338 1 DCUP_AQUAE

ALIGNMENTS

RESULT 1
VGLF_MEASY STANDARD; PRT; 534 AA.
AC P26032;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion Glycoprotein F1].
GN F.
OS Measles virus (strain Yamagata-1) (Subacute sclerosing panencephalitis
OS virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID:11239;
RN [1]_TaxID:11239;
RP SEQUENCE FROM N.A.
RX MEDLINE=90385702; PubMed=1698327;
RA Komase K., Haga T., Yoshikawa Y., Sato T.A., Yamanouchi K.;
RT "Molecular analysis of structural protein genes of the Yamagata-1
RT strain of defective subacute sclerosing panencephalitis virus. IV.
RT Nucleotide sequence of the fusion gene.";
RL Virus Genes 4:173-181(1990).
CC -!- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; D10548; BAA01405.1; -
CC HSSP; P04849; LSVP.
DR InterPro; IPR000776; Fusion gly.
DR Pfam; PF00523; fusion gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 534 FUSION GLYCOPROTEIN FO.
FT CHAIN 24 112 PROTEIN F2.
FT CHAIN 113 534 PROTEIN F1.
FT TRANSMEM 113 136 POTENTIAL.
FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 495 515 POTENTIAL.
FT DOMAIN 516 534 CYTOPLASMIC (POTENTIAL).
FT DISULFID 58 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE 534 AA; 57963 MW; F5B2175B643844D CRC64;

Qy 3 LSEIKGVIVHRLEGV 17

AC P08300;

DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Measles virus (strain Edmonston) (Subacute sclerosing panencephalitis
OS virus),
OS Measles virus (strain Halle) (Subacute sclerosing panencephalitis
OS virus),
OS Measles virus (strain Leningrad-16) (Subacute sclerosing panencephalitis
OS virus),
OS Measles virus (strain Edmonston-Zagreb) (Subacute sclerosing
OS panencephalitis virus),
OS Measles virus (strain Philadelphia-26) (Subacute sclerosing
OS panencephalitis virus), and
OS Measles virus (strain Edmonston B) (Subacute sclerosing panencephalitis
OS virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
OX NCBI_TaxID=11235, 11236, 70147, 70148, 70149, 70148, 70146;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Edmonston;
RX MEDLINE=87071568; PubMed=3788062;
RA Richardson C.D., Hull D., Greer P., Hasel K., Berkovich A.,
RA Englund G., Bellini W.J., Rima B., Lazzarini R.A.;
RT "The nucleotide sequence of the mRNA encoding the fusion protein of
RT measles virus (Edmonston strain): a comparison of fusion proteins
RT from several different paramyxoviruses.";
RL Virology 155:508-523 (1986).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Halle;
RX MEDLINE=87224816; PubMed=3585281;
RA Buckland R., Gerald C., Barker R., Wild T.F.;
RT "Fusion glycoprotein of measles virus: nucleotide sequence of the
RT gene and comparison with other paramyxoviruses.";
RN J. Gen. Virol. 68:1695-1703 (1987).
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=Edmonston;
RX MEDLINE=90085790; PubMed=2596022;
RA Cattaneo R., Schmid A., Spielhofer P., Kaelin K., Bacsko K.,
RA Meulen V., Pardowitz J., Flanagan S., Rima B.K., Udem S.A.;
RT "Mutated and hypermutated genes of persistent measles viruses which
RT caused lethal human brain diseases.";
RL Virology 173:415-425 (1989).
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=Edmonston;
RX MEDLINE=92263801; PubMed=1585653;
RA Schmid A., Spielhofer P., Cattaneo R., Bacsko K., Ter Meulen V.,
RA Billeter M.A.;
RT "Subacute sclerosing panencephalitis is typically characterized by
RT alterations in the fusion protein cytoplasmic domain of the
RT persisting measles virus.";
RL Virology 188:910-915 (1992).
RN [5]
RN SEQUENCE FROM N.A.
RP STRAIN=Edmonston, Leningrad-16, and Edmonston-Zagreb;
RX MEDLINE=94249283; PubMed=8191786;
RA Rota J.S., Wang Z.D., Rota P.A., Bellini W.J.;
RT "Comparison of sequences of the H, F, and N coding genes of measles
RT virus vaccine strains.";
RL Virus Res. 31:317-330 (1994).
RN [6]
RN SEQUENCE FROM N.A.
RP STRAIN=Philadelphia-26;
RX MEDLINE=94303181; PubMed=8030232;
RA Hummel K.B., Vanchiere J.A., Bellini W.J.;
RT "Restriction of fusion protein mRNA as a mechanism of measles virus
RT persistence.";

RL Virology 202:665-672 (1994).
RN [7]
RN SEQUENCE FROM N.A.
RP STRAIN=Edmonston B;
RA Billeter M.A.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC 2- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES P-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC 3- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14915; AAA46423.1; -
DR EMBL; X05597; CAA29090.1; ALT_INIT.
DR EMBL; K01711; AAA75498.1; ALT_INIT.
DR EMBL; K01711; AAA75499.1; -
DR EMBL; U03657; AAA56647.1; ALT_INIT.
DR EMBL; U03659; AAA56649.1; ALT_INIT.
DR EMBL; U03670; AAA56660.1; ALT_INIT.
DR EMBL; U08416; AAA50550.1; ALT_INIT.
DR EMBL; Z66517; CAA31367.1; ALT_INIT.
DR EMBL; Z66517; CAA31368.1; -
DR HSP; P04849; LSVE.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 550 FUSION GLYCOPROTEIN F0.
FT CHAIN 24 112 PROTEIN F2.
FT CHAIN 113 550 PROTEIN F1.
FT TRANSMEM 113 136 POTENTIAL.
FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 495 515 POTENTIAL.
FT DOMAIN 516 550 CYTOPLASMIC (POTENTIAL).
FT DISULFID 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 550 AA; 59532 MW; 7AA4F1CA82169093 CRC64;
Query Match 45.0%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 LSEIKGVIVHRLGV 17
Db 288 LSEIKGVIVHRLGV 302
RESULT 5
VGLF_RINDR STANDARD; PRT; 546 AA.
AC P41356;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain R80K) (RDV).
OS Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=36409;
RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=9508609; PubMed=7996154;
RA Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;
RT "Nucleotide sequence comparisons of the fusion protein gene from
RT virulent and attenuated strains of rinderpest virus.";
RL J. Gen. Virol. 75:3611-3617(1994).
CC -!- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z30700; CAA83186.1; -.
DR EMBL; Z30697; CAA83181.1; -.
DR PIR; S47305; S47305.
DR HSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58705 MW; ED3DF8AFDECB95 CRC64;

Query Match 44.4%; Score 71; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.0087;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGV 17
DB 284 LSEIKGVIVHRLGV 298

RESULT 6
VGLF_RINDL STANDARD; PRT; 546 AA.
AC P10864;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain L) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11243;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8219541; PubMed=3285575;
RA Tsukiyama K., Yoshikawa Y., Yamanouchi K.;
RT "Fusion glycoprotein (F) of rinderpest virus: entire nucleotide
RT sequence of the F mRNA, and several features of the F protein.";
RL Virology 164:523-530(1998).

```

```

CC -!- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M20870; AAA47399.1; -.
DR PIR; A28921; VGNZRL.
DR HSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 546 AA; 985029418F28FFB5 CRC64;

Query Match 41.2%; Score 66; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.048;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGV 17
DB 284 LSEIKGVIVHRLGV 298

RESULT 7
VGLF_RINDL STANDARD; PRT; 546 AA.
AC P12574;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain Kabete O) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11243;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89322864; PubMed=3413983;
RA Hsu D., Yamanaka M., Miller J., Dale B., Grubman M., Yilma T.;
RT "Cloning of the fusion gene of rinderpest virus: comparative sequence
RT analysis with other morbilliviruses.";
RL Virology 166:149-153(1988).
CC -!- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M21514; AAA47400.1; -;
 CC PIR; A31051; VGNZK.

CC HSP; P04849; 1SVF.
 CC InterPro; IPR000776; Fusion_gly; 1.

CC Pfam; PF00523; Fusion_gly; 1.
 CC Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.

FT SIGNAL 1 19
 FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.

FT CHAIN 20 108 F2 PROTEIN.
 FT CHAIN 109 546 F1 PROTEIN.

FT DOMAIN 104 108 ARG-RICH (BASIC).
 FT TRANSMEM 109 133 POTENTIAL.

FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
 FT TRANSMEM 484 513 POTENTIAL.

FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
 FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 546 AA; 58662 MW; 47674DCC18BCFCF CRC64;

Query Match 40.6%; Score 65; DB 1; Length 546;
 Best Local Similarity 86.7%; Pred. No. 0.068; Mismatches 1; Indels 0; Gaps 0;
 Matches 13; Conservative 1;

QY 3 LSEIKGVIVHRLEGV 17
 |||||:|||||

DB 284 LSEIKGVIIHLEGV 298
 |||||:|||||

RESULT 8
 VGLF CDVO STANDARD; PRT; 662 AA.

AC P12569; O65991;
 DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
 DE Fusion glycoprotein F1].

OS F.
 OS Canine distemper virus (strain Onderstepoort) (CDV).

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.

OX NCBI_TaxID=11233;
 [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=88129050; PubMed=3433924;

RA Barrett T., Clarke D.K., Evans S.A., Rima B.K.;
 RA "The nucleotide sequence of the gene encoding the F protein of canine

RT distemper virus: a comparison of the deduced amino acid sequence with
 RT other paramyxoviruses.";

RL Virus Res. 8:373-386(1987).
 [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=93227696; PubMed=8470428;

RA Wild T.F., Bernard A., Spohner D., Villevall D., Drillion R.;
 RA "Vaccination of mice against canine distemper virus-induced

RT encephalitis with vaccinia virus recombinants encoding measles or
 RT canine distemper virus antigens.";

RL Vaccine 11:438-444(1993).
 CC -!- FUNCTION: This protein directs fusion of viral and cellular

CC membranes.
 CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2

CC -!- LINKED BY A DISULFIDE BOND.
 CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein

CC family.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

```
CC family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X16566; CAA34567.1; -.
CC DR EMBL; X16566; CAA34568.1; ALT_INIT.
CC DR HSSP; P04849; ISVF.
CC DR InterPro; IPR000776; Fusion_gly.
CC DR Pfam; PF00523; Fusion_gly; 1.
CC KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
CC FT SIGNAL 1 26 BY SIMILARITY.
CC FT CHAIN 27 529 FUSION GLYCOPROTEIN F0.
CC FT CHAIN 116 529 PROTEIN F2.
CC FT CHAIN 116 529 POTENTIAL.
CC FT TRANSMEM 116 139 EXTRACELLULAR (POTENTIAL).
CC FT DOMAIN 140 497 POTENTIAL.
CC FT TRANSMEM 498 518 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 519 529 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
CC FT DISULFID 71 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 529 AA; 57331 MW; AE987BC9P07E9AA9 CRC64;
Query Match 40.0%; Score 64; DB 1; Length 529;
Best Local Similarity 93.3%; Pred. No. 0.093;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 LSEIKGVIVHRLGV 17
Db 291 LSEIKGVIVHRLGV 305
RESULT 10
VGLF PHODV STANDARD; PRT; 631 AA.
AC P2886;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Phocine distemper virus (PDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
ON NCBI_TaxID=11240;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate DK88-4A;
RX MEDLINE=92113538; PubMed=1765768;
RA Koevanees J., Blixenkron-Moeller M., Sharma B., Oervell C.,
RA Norby E.;
RT "The nucleotide sequence and deduced amino acid composition of the
RT haemagglutinin and fusion proteins of the morbillivirus phocid
RT distemper virus.";
RT J. Gen. Virol. 72:2959-2966(1991).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Ulster/88;
RC STRAIN=92398437; PubMed=1524494;
RA Curran M.D., Lu Y.J., Rima B.K.;
RX "The fusion protein gene of phocine distemper virus: nucleotide and
RT deduced amino acid sequences and a comparison of morbillivirus fusion
RT proteins.";
RL Arch. Virol. 126:159-169(1992).
RN [3]
```

```
RP SEQUENCE OF 95-631 FROM N.A.
RC STRAIN=Ulster/88; PubMed=2264246;
RX MEDLINE=91089508; PubMed=2264246;
RA Curran M.D., Loan D.O., Rima B.K., Kennedy S.;
RT "Nucleotide sequence analysis of phocine distemper virus reveals its
RT distinctness from canine distemper virus.";
RL Vet. Rec. 127:430-431(1990).
CC -1- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D10371; BAA01206.1; -.
CC DR EMBL; A48346; A48346.
CC DR PIR; J01368; VGNZPD.
CC DR HSSP; P04849; ISVF.
CC DR InterPro; IPR000776; Fusion_gly.
CC DR Pfam; PF00523; fusion_gly; 1.
CC KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
CC FT SIGNAL 1 631 FUSION GLYCOPROTEIN F0.
CC FT CHAIN 2 186 F2 PROTEIN.
CC FT CHAIN 194 631 F1 PROTEIN.
CC FT DISULFID 149 276 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
CC FT TRANSMEM 89 106 POTENTIAL.
CC FT TRANSMEM 194 212 POTENTIAL.
CC FT TRANSMEM 575 595 POTENTIAL.
CC FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CONFLICT 63 63 I -> V (IN REF. 2).
CC SQ SEQUENCE 631 AA; 68873 MW; D1FC87CDD426E9B8 CRC64;
Query Match 40.0%; Score 64; DB 1; Length 631;
Best Local Similarity 80.0%; Pred. No. 0.11;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 LSEIKGVIVHRLGV 17
Db 369 LSEIKGVIVHRLGV 383
RESULT 11
GONI_PIG STANDARD; PRT; 91 AA.
AC P49921;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadolibarin I precursor [Contains: Gonadolibarin I (LH-RH I)
DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I].
GN GNRH1 OR GNRH.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RA Weener G.D., Matteri R.L., Becker B.A.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
```


DR GO; GO:0005625; C:soluble fraction; TAS.
DR GO; GO:0005183; F:lutinizing hormone-releasing factor activity; TAS.
DR GO; GO:0007267; P:cell-cell signaling; TAS.
DR GO; GO:0007275; P:development; TAS.
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pharmacological; Signal; Polymorphism;
KW Pyroglutamate carboxylic acid.
FT SIGNAL 1 23 PROGNADOLIBERIN I.
FT CHAIN 24 92 GONADOLIBERIN I.
FT PEPTIDE 24 33 GnRH-ASSOCIATED PEPTIDE I.
FT PEPTIDE 37 92 GnRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT VARIANT 16 16 W->S (in dbSNP:6185).
FT FTID=VAR 013943.
SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;

Query Match 35.9%; Score 57.5; DB 1; Length 92;
Best Local Similarity 80.0%; Pred. No. 0.14; Mismatches 0; Indels 1; Gaps 1;
Matches 12; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 17 VEG-PSLHWSYGLRP 30
DB 18 VEGCSQHSYGLRP 32
||| |||||
RESULT 13
GONI_MOUSE STANDARD; PRT; 90 AA.
ID GONI_MOUSE
AC P13562;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Prognadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)]
DE (lutinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor I).
DE GNRH1 OR GNRH.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87069928; PubMed=3024317;
RA Mason A.J., Hayflick J.S., Zeller R.T., Young W.S. III,
Phillips H.S., Nikolic K., Seeburg P.H.;
RT "A deletion truncating the gonadotropin-releasing hormone gene is responsible for hypogonadism in the hpg mouse.";
RL Science 234:1366-1371(1986).
RC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates the secretion of both lutinizing and follicle-stimulating hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC

EMBL; M14872; AAA37717.1; --
PIR; A47578; RHMSG.
MGD; MG1:95789; GnRH.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyroglutamate carboxylic acid.
FT SIGNAL 1 21 PROGNADOLIBERIN I.
FT CHAIN 22 90 GONADOLIBERIN I.
FT PEPTIDE 22 31 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT PEPTIDE 35 90 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.
FT ACT_SITE 24 24
FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
SQ SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;

Query Match 35.6%; Score 57; DB 1; Length 90;
Best Local Similarity 70.6%; Pred. No. 0.17;
Matches 12; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 14 LEGVEGSPSLHWSYGLRP 30
DB 16 LEGCS--SQHWSYGLRP 30
||| |||||
RESULT 14
GONI_RAT STANDARD; PRT; 92 AA.
ID GONI_RAT
AC P07490;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Prognadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)]
DE (lutinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor I).
DE GNRH1 OR GNRH.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and prolactin release-inhibiting factor in human and rat.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89384661; PubMed=2476669;
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT "The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic expression.";
RL Mol. Endocrinol. 3:1257-1262(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Thymus;
CC MEDLINE=93105480; PubMed=468115;
RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic lutinizing hormone-releasing hormone mRNA.";
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN [4]
RP SEQUENCE OF 1-47 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=87149087; PubMed=3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;

```

RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus.";
RL Science 235:1514-1517(1987).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Central nervous system.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S50870; AAB24572.1; -.
DR EMBL; M12579; AAA41263.1; -.
DR EMBL; M31570; AAA41264.1; -.
DR EMBL; M15527; AAA42141.1; ALT_SEQ.
DR EMBL; M15529; AAA42139.1; -.
DR EMBL; M15528; -. NOT ANNOTATED CDS.
DR PIR; A40147; RHRTS.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92 PROGONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT ACT_SITE 26 26 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACT_SITE 26 26 ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;

Query Match 35.6%; Score 57; DB 1; Length 92;
Best Local Similarity 70.6%; Pred. No. 0.17; 3; Indels 2; Gaps 1;
Matches 12; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 14 LEGVEGPSLHWSYGLRP 30
Db 18 LEGCS--SQHWSYGLRP 32
||| |||||
18 LEGCS--SQHWSYGLRP 32

RESULT 15
GONI_MACMU
ID GONI_MACMU STANDARD; PRT; 67 AA.
AC P55247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment).
DE GN RH1 OR GN RH OR LHRH.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Hypothalamus;
RX MEDLINE=95124501; PubMed=7545971;
RA Ma Y.J.; Costa M.E.; Ojeda S.R.;

```

```

RT "Developmental expression of the genes encoding transforming growth
RT factor alpha and its receptor in the hypothalamus of female rhesus
RT macaques.";
RL Neuroendocrinology 60:346-359(1994).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S75918; AAB33096.1; -.
DR PIR; I78541; I78541.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Pyrrolidone carboxylic acid.
FT NON_TER 1 1
FT SIGNAL <1 5 BY SIMILARITY.
FT CHAIN 6 >67 PROGONADOLIBERIN I.
FT PEPTIDE 6 15 GONADOLIBERIN I.
FT PEPTIDE 19 >67 GNHR-ASSOCIATED PEPTIDE I.
FT ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACT_SITE 8 8 ACTIVITY (BY SIMILARITY).
FT MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 15 15 SIMILARITY).
FT MOD_RES 67 67 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
FT NON_TER 67 67 SIMILARITY).
SQ SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

Query Match 33.8%; Score 54; DB 1; Length 67;
Best Local Similarity 90.0%; Pred. No. 0.35;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 SLHWSYGLRP 30
Db 5 SQHWSYGLRP 14
||| |||||
5 SQHWSYGLRP 14

Search completed: March 10, 2004, 09:13:52
Job time : 5.66926 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 30.6381 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-9
Perfect score: 160
Sequence: 1 KLLSEIKGVIVHRLEGVSPSLHWSYGLRPX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phages:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 72 | 45.0 | 534 | 12 Q04243 | Q04243 measles vir |
| 2 | 72 | 45.0 | 537 | 12 Q04242 | Q04242 measles vir |
| 3 | 72 | 45.0 | 545 | 12 Q9PX44 | Q9PX44 measles vir |
| 4 | 72 | 45.0 | 546 | 12 Q91HA5 | Q91HA5 rinderpest |
| 5 | 72 | 45.0 | 550 | 12 P90331 | P90331 measles vir |
| 6 | 72 | 45.0 | 550 | 12 Q9QEX0 | Q9QEX0 measles vir |
| 7 | 72 | 45.0 | 550 | 12 Q9QEW9 | Q9QEW9 measles vir |
| 8 | 72 | 45.0 | 550 | 12 P90330 | P90330 measles vir |
| 9 | 72 | 45.0 | 550 | 12 Q9QEW7 | Q9QEW7 measles vir |
| 10 | 72 | 45.0 | 550 | 12 Q9WMK4 | Q9WMK4 measles vir |
| 11 | 72 | 45.0 | 550 | 12 Q89435 | Q89435 measles vir |
| 12 | 72 | 45.0 | 550 | 12 Q8V049 | Q8V049 measles vir |
| 13 | 72 | 45.0 | 550 | 12 Q9YJ94 | Q9YJ94 measles vir |
| 14 | 72 | 45.0 | 550 | 12 Q9QEX1 | Q9QEX1 measles vir |
| 15 | 72 | 45.0 | 550 | 12 Q9QEW8 | Q9QEW8 measles vir |
| 16 | 72 | 45.0 | 553 | 12 Q93055 | Q93055 measles vir |

Q9IC36 measles vir
P88973 measles vir
Q83536 measles vir
O11383 measles vir
Q91FK2 measles vir
Q83533 measles vir
Q83525 measles vir
Q83518 measles vir
P88974 measles vir
Q83527 measles vir
Q83521 measles vir
Q83530 measles vir
Q91248 measles vir
Q91QP2 measles vir
Q04244 measles vir
Q9PWU4 measles vir
Q9QEW6 measles vir
O11380 measles vir
O66147 cetacean mo
Q9YJ99 canine dist
Q8QV06 canine dist
Q9DX22 canine dist
Q91KN3 canine dist
Q9YKL7 canine dist
Q89327 canine dist
O66409 dolphin mor
Q56852 dolphin mor
Q83629 measles vir
Q84926 peste-des-p

ALIGNMENTS

RESULT 1

Q04243 PRELIMINARY; PRT; 534 AA.
AC Q04243;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Fusion protein.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses;
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89003063; PubMed=3167982;
RA Cattaneo R., Schmid A., Eschle D., Baczko K., ter Meulen V.,
RA Billeter M.A.;
RT "Biased hypermutation and other genetic changes in defective measles
RT viruses in human brain infections.";
RL Cell 55:255-265(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Cattaneo R., Billeter M.A.;
RL Virology 0:0-0(0).
DR EMBL; X16568; CAA34581.1; -.
DR EMBL; X16568; CAA34582.1; -.
DR HSSP; P04849; 1SVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion Gly.
DR Pfam; PF00523; fusion gly; 1.
SQ SEQUENCE 534 AA; 57899 MW; 637245323B5B5044 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.073; 0; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;
QY 3 LSEIKGVIVHRLEGV 17

```

Db      291 LSEIKGVVHRLEGV 305
|||||
RESULT 2
Q04242      PRELIMINARY;      PRT;      537 AA.
AC Q04242;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Fusion protein.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89003063; PubMed=3167982;
RA Cattaneo R., Schmid A., Eschle D., Bacsko K., ter Meulen V.,
RA Billette M.A.;
RT "Biased hypermutation and other genetic changes in defective measles
RT viruses in human brain infections.";
RL Cell 55:255-265 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Cattaneo R., Billette M.A.;
RL Virology 0:0-0(0).
DR EMBL; X16567; CAA34574.1; --
DR EMBL; X16567; CAA34575.1; --
DR HSP; P04849; ISVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; I.
SQ SEQUENCE 537 AA; 58275 MW; D0A60AC66D979E06 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 537;
Best Local Similarity 100.0%; Pred. No. 0.073;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 LSEIKGVVHRLEGV 17
|||||
Db      291 LSEIKGVVHRLEGV 305
|||||

RESULT 3
Q9PXA4      PRELIMINARY;      PRT;      545 AA.
AC Q9PXA4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OSA-3;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179440; AAF02705.1; --
DR EMBL; AF179439; AAF02704.1; --
DR HSP; P04849; ISVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.

```

```

DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; I.
SQ SEQUENCE 545 AA; 58907 MW; 0234C28AE193E77D CRC64;

Query Match 45.0%; Score 72; DB 12; Length 545;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 LSEIKGVVHRLEGV 17
|||||
Db      288 LSEIKGVVHRLEGV 302
|||||

RESULT 4
Q91HA5      PRELIMINARY;      PRT;      546 AA.
AC Q91HA5;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Fusion protein.
GN F.
OS Rinderpest virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11241;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K;
RC MEDLINE=21014265; PubMed=1186456;
RA Aianot P.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,
RA Gusev A.A.;
RT "Primary structure of the F-gene from Rinderpest virus strain K.";
RL Mol. Gen. Microbiol. Virusol. 4:29-33 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K;
RA Ayanot P.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,
RA Gusev A.A.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY035887; AAK63190.1; --
DR PIR; PQ0866; PQ0866.
DR PIR; PQ0867; PQ0867.
DR PIR; PQ0873; PQ0873.
DR GO; GO:0019039; P:Viral-induced cell-cell fusion; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; I.
SQ SEQUENCE 546 AA; 58572 MW; 449B2B2DD7405F0B CRC64;

Query Match 45.0%; Score 72; DB 12; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 LSEIKGVVHRLEGV 17
|||||
Db      284 LSEIKGVVHRLEGV 298
|||||

RESULT 5
P90331      PRELIMINARY;      PRT;      550 AA.
ID P90331
AC P90331;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Fusion protein.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;

```

```

RN  SEQUENCE FROM N.A.
RC  STRAIN=NAGAHATA;
RA  Sheng J., Watanabe M., Ueda S.;
RT  "Selection of a neurotropic variant of measles virus.";
RL  Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=NAGAHATA;
RA  Sheng J., Nakanishi M., Watanabe M., Ueda S.;
RT  "An amino acid alteration of F protein responsible for the enhanced
RL  fusogenicity of measles virus.";
RN  Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN=NAGAHATA;
RA  Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RT  Ogura H.;
RN  "Nucleotide sequences of the fusion protein gene of subacute
RT  sclerosing panencephalitis viruses; deduced amino acid sequences
RT  showed the cytoplasmic domain highly mutated --truncated, elongated or
RT  predicted secondary structure changed.";
RL  Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL; D63926; BRA0958.1; -.
DR  EMBL; AF179431; AAF02696.1; -.
DR  PIR; PQ0376; PQ0376.
DR  HSSP; P04849; 1SVF.
DR  GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR  GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR  InterPro; IPR000776; Fusion_gly.
DR  Pfam; PF00523; fusion_gly; 1.
SQ  SEQUENCE 550 AA; 59530 MW; 97C991C7E2169839 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
DB 288 LSEIKGVIVHRLEGV 302

RESULT 6
QOEX0 PRELIMINARY; PRT; 550 AA.
AC QOEX0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Toyoshima;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses; deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179432; AAF02697.1; -.
DR PIR; PQ0376; PQ0376.
DR HSSP; P04849; 1SVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59504 MW; 2AA969D37FA5CA17 CRC64;

```

```

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
DB 288 LSEIKGVIVHRLEGV 302

RESULT 7
QOEW9 PRELIMINARY; PRT; 550 AA.
AC QOEW9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OSA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses; deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179436; AAF02701.1; -.
DR PIR; PQ0376; PQ0376.
DR HSSP; P04849; 1SVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59405 MW; 0AE6DBFCDD22BBA CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
DB 288 LSEIKGVIVHRLEGV 302

RESULT 8
P90330 PRELIMINARY; PRT; 550 AA.
ID P90330;
AC P90330;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fusion protein.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagahata(HB);
RA Sheng J., Watanabe M., Ueda S.;
RT "Selection of a neurotropic variant of measles virus.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagahata(HB);
RA Sheng J., Nakanishi M., Watanabe M., Ueda S.;
RT "An amino acid alteration of F protein responsible for the enhanced

```

```

RT fusogenicity of measles virus."
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; D63924; BAA09951.1; -.
DR PIR; PQ0376; PQ0376.
DR HSP; P04849; LSVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR CO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59589 MW; 7387BD457ABA39B7 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
DB 288 LSEIKGVIVHRLEGV 302

RESULT 9
Q9QEW7 PRELIMINARY; PRT; 550 AA.
AC Q9QEW7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
CX NCBI_TaxID=11234;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=OSA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RA "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: Deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179438; AA02703.1; -.
DR PIR; PQ0376; PQ0376.
DR HSP; P04849; LSVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR CO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59333 MW; 086B51FED5582BBA CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
DB 288 LSEIKGVIVHRLEGV 302

RESULT 10
Q9WMK4 PRELIMINARY; PRT; 550 AA.
AC Q9WMK4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
CX NCBI_TaxID=11234;

```

```

[1]
RN SEQUENCE FROM N.A.
RC STRAIN=WTF;
RX MEDLINE=99329215; PubMed=10400788;
RA Johnston I.C., Ter Meulen V., Schneider-Schaulies J.,
RA Schneider-Schaulies S.;
RT "A recombinant measles vaccine virus expressing wild-type
RT glycoproteins: consequences for viral spread and cell tropism.";
RL J. Virol. 73:6903-6915(1999).
DR EMBL; AJ133108; CAB38075.1; -.
DR PIR; PQ0376; PQ0376.
DR HSP; P04849; LSVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR CO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59580 MW; 825549996885D862 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
DB 288 LSEIKGVIVHRLEGV 302

RESULT 11
Q89495 PRELIMINARY; PRT; 550 AA.
AC Q89495;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
CX NCBI_TaxID=11234;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=92230209; PubMed=1566568;
RA Rota J.S., Hummel K.B., Rota P.A., Bellini W.J.;
RT "Genetic variability of the glycoprotein genes of current wild-type
RT measles isolates.";
RL Virology 188:135-142(1992).
DR EMBL; M81903; AAA46422.1; -.
DR EMBL; M81901; AAA46421.1; -.
DR PIR; PQ0376; PQ0376.
DR HSP; P04849; LSVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR CO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59564 MW; A78EEC9CD6268E58 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
DB 288 LSEIKGVIVHRLEGV 302

RESULT 12
Q8V049 PRELIMINARY; PRT; 550 AA.
AC Q8V049;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

```

DE F. Fusion protein.
GN Measles virus.
OS Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OC NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=G954;
RC MEDLINE=21635526; PubMed=1173423;
RA Waku Koumou D., Wild T.F.;
RL "Adaptation of wild-type measles virus to tissue culture.";
RT J. Virol. 76:1505-1509(2002).
DR EMBL; AY059392; AAL29688.1; -.
DR PIR; PQ0376; PQ0376
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; I.
SQ SEQUENCE 550 AA; 59551 MW; 9A7A4BA99E4DA8E9 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVVHRLEGV 17
Db |||||
288 LSEIKGVVHRLEGV 302

RESULT 13
Q9YJU94 PRELIMINARY; PRT; 550 AA.
ID Q9YJU94
AC Q9YJU94
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OC NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98440529; PubMed=9765410;
RA Takeda M., Kato A., Kobune F., Sakata H., Li Y., Shioda T., Sakai Y.,
RA Asakawa M., Nagai Y.;
RL "Measles virus attenuation associated with transcriptional impediment
RT and a few amino acid changes in the polymerase and accessory
RT proteins.";
RL J. Virol. 72:8690-8696(1998).
DR EMBL; AB012949; BAA33877.1; -.
DR EMBL; AB012948; BAA33871.1; -.
DR PIR; PQ0376; PQ0376.
DR HSSP; P04849; 1SVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; I.
SQ SEQUENCE 550 AA; 59512 MW; 7AA4F1D117197BF9 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVVHRLEGV 17
Db |||||
288 LSEIKGVVHRLEGV 302

RESULT 14
Q9QEX1 PRELIMINARY; PRT; 550 AA.
ID Q9QEX1
AC Q9QEX1
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OC NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Masusako;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RL "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses; deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179430; AAF02695.1; -.
DR PIR; PQ0376; PQ0376.
DR HSSP; P04849; 1SVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; I.
SQ SEQUENCE 550 AA; 59559 MW; 609EE024A7E59C54 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVVHRLEGV 17
Db |||||
288 LSEIKGVVHRLEGV 302

RESULT 15
Q9QEW8 PRELIMINARY; PRT; 550 AA.
ID Q9QEW8
AC Q9QEW8
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OC NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=OSA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RL "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses; deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179437; AAF02702.1; -.
DR PIR; PQ0376; PQ0376.
DR HSSP; P04849; 1SVF.
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; I.
SQ SEQUENCE 550 AA; 59315 MW; 086E51FED235EBBA CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


Qy 3 LSEIKGVIVHRLEGV 17
|||
Db 288 LSEIKGVIVHRLEGV 302

Search completed: March 10, 2004, 09:25:31
Job time : 32.6381 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:48 ; Search time 46.6809 Seconds
(without alignments)
187.635 Million cell updates/sec

Title: US-09-848-834A-9

Perfect score: 160

Sequence: 1 KLLSEIKGVIVHRLEGVPSLHWSVGLRPX 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Geneseq_29Jan04:*
- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|------------|--------------------|
| 1 | 159 | 99.4 | 31 | 5 AAU1420 | Aau1420 Synthetic |
| 2 | 159 | 99.4 | 47 | 5 AAU1428 | Aau1428 Synthetic |
| 3 | 117 | 73.1 | 25 | 2 AAR62705 | Aar62705 LHRH-cont |
| 4 | 117 | 73.1 | 42 | 2 AAR62708 | Aar62708 LHRH-cont |
| 5 | 116 | 72.5 | 27 | 2 AAR62707 | Aar62707 LHRH-cont |
| 6 | 116 | 72.5 | 27 | 3 AAY68567 | Aay68567 Peptide i |
| 7 | 116 | 72.5 | 27 | 3 AAY91156 | Aay91156 MVF Th ep |
| 8 | 116 | 72.5 | 45 | 2 AAR62721 | Aar62721 LHRH-cont |
| 9 | 116 | 72.5 | 45 | 7 ADD89949 | Add89949 LHRH-pept |
| 10 | 113 | 70.6 | 27 | 3 AAY91163 | Aay91163 Modified |
| 11 | 110 | 68.8 | 31 | 3 AAY91175 | Aay91175 Modified |
| 12 | 108 | 67.5 | 27 | 3 AAY91161 | Aay91161 Modified |
| 13 | 108 | 67.5 | 27 | 3 AAY91167 | Aay91167 Modified |
| 14 | 108 | 67.5 | 45 | 3 AAY68573 | Aay68573 Peptide i |
| 15 | 108 | 67.5 | 45 | 3 AAY91165 | Aay91165 Modified |
| 16 | 107 | 66.9 | 31 | 3 AAY91179 | Aay91179 Modified |
| 17 | 106 | 66.2 | 28 | 3 AAY91158 | Aay91158 Modified |
| 18 | 105 | 65.6 | 31 | 3 AAY68582 | Aay68582 Peptide i |
| 19 | 105 | 65.6 | 31 | 3 AAY91173 | Aay91173 Modified |
| 20 | 105 | 65.6 | 47 | 3 AAY68583 | Aay68583 Peptide i |
| 21 | 105 | 65.6 | 47 | 3 AAY91180 | Aay91180 Inv epit |
| 22 | 105 | 65.6 | 49 | 3 AAY91177 | Aay91177 Modified |
| 23 | 103 | 64.4 | 27 | 3 AAY68575 | Aay68575 Peptide i |
| 24 | 103 | 64.4 | 27 | 3 AAY91170 | Aay91170 Modified |
| 25 | 103 | 64.4 | 35 | 3 AAY91242 | Aay91242 Modified |

| | | | | | |
|----|-----|------|----|------------|--------------------|
| 26 | 100 | 62.5 | 28 | 2 AAR62726 | Aar62726 LHRH-cont |
| 27 | 100 | 62.5 | 46 | 2 AAR62728 | Aar62728 LHRH-cont |
| 28 | 100 | 62.5 | 47 | 3 AAY68586 | Aay68586 Peptide i |
| 29 | 100 | 62.5 | 47 | 3 AAY91183 | Aay91183 Inv epit |
| 30 | 99 | 61.9 | 28 | 3 AAY91159 | Aay91159 Modified |
| 31 | 99 | 61.9 | 75 | 6 ABP72236 | Abp72236 Frizzled |
| 32 | 94 | 58.8 | 28 | 3 AAY91157 | Aay91157 Modified |
| 33 | 89 | 55.6 | 27 | 2 AAR74261 | Aar74261 SSALL TH1 |
| 34 | 87 | 54.4 | 34 | 2 AAU05619 | Aau05619 MVFTH-GG |
| 35 | 86 | 53.8 | 34 | 5 AAU1424 | Aau1424 Synthetic |
| 36 | 84 | 52.5 | 27 | 3 AAY68576 | Aay68576 Peptide i |
| 37 | 84 | 52.5 | 27 | 3 AAY91171 | Aay91171 Modified |
| 38 | 84 | 52.5 | 45 | 3 AAY68577 | Aay68577 Peptide i |
| 39 | 84 | 52.5 | 45 | 3 AAY91172 | Aay91172 Modified |
| 40 | 81 | 50.6 | 31 | 3 AAY68587 | Aay68587 Peptide i |
| 41 | 81 | 50.6 | 31 | 3 AAY91184 | Aay91184 Modified |
| 42 | 79 | 49.4 | 27 | 3 AAY68574 | Aay68574 Peptide i |
| 43 | 79 | 49.4 | 27 | 3 AAY91168 | Aay91168 Modified |
| 44 | 79 | 49.4 | 27 | 3 AAY91164 | Aay91164 Modified |
| 45 | 79 | 49.4 | 27 | 3 AAY91169 | Aay91169 Modified |

ALIGNMENTS

RESULT 1

AAU1420

ID AAU1420 standard; peptide; 31 AA.

XX

AC AAU1420;

XX

DT 12-MAR-2002 (first entry)

XX

DE Synthetic immunogen peptide 1.

XX

KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;

KW luteinising hormone releasing hormone; LHRH; contraceptive;

KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;

KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;

KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX Measles virus.

OS Mammalia.

OS Synthetic.

OS Chimeric.

XX

Key Location/Qualifiers

FT Peptide 1..18

FT Peptide /note= "Measles virus fusion protein F epitope"

FT Peptide 19..22

FT Peptide /note= "Spacer peptide"

FT Peptide 23..31

FT Peptide /note= "Gonadotropin releasing hormone epitope"

FT Modified-site 31

FT /note= "Amidated glycine or glycinamide"

XX

WO200185763-A2.

PN

XX

PD 15-NOV-2001.

XX

PF 04-MAY-2001; 2001WO-US014363.

XX

PR 05-MAY-2000; 2000US-0202328P.

XX

XX (APHT-) APHTON CORP.

XX

PA Grimes S, Michaeli D, Stevens VC;

XX

PI WPI; 2002-049440/06.

XX

DR Novel synthetic immunogen for inducing immune response against

PT gonadotropin releasing hormone, comprises fusion peptide having

PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope

PT

PT or its analog.
 XX
 PS Claim 11; Page 7; 43pp; English.
 XX
 CC The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX
 XX Sequence 31 AA;
 SQ
 Query Match 99.4%; Score 159; DB 5; Length 31;
 Best Local Similarity 100.0%; Pred. NO. 9.3e-17; Mismatches 0; Indels 0; Gaps 0;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KLLSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30
 |||||
 DB 1 KLLSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30
 |||||
 RESULT 2
 AAU11428
 ID AAU11428 standard; peptide; 47 AA.
 AC AAU11428;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE Synthetic immunogen peptide 9.
 XX
 KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
 XX
 OS Plasmodium falciparum.
 OS Mammalia.
 OS Synthetic.
 OS Chimeric.
 XX
 PH Key Location/Qualifiers
 FT Peptide 1..10
 FT /note= "Gonadotropin releasing hormone epitope (1..10
 FT aa)"
 FT Misc-difference 1
 FT /label= OTHER
 FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
 FT OS
 FT Peptide 11..16
 FT /note= "Spacer peptide"
 FT FT Peptide 17..34
 FT /note= "Malaria CSP protein (288-302 aa)"
 FT FT Peptide 35..38
 FT /note= "Spacer peptide"
 FT FT Peptide 39..47
 FT /note= "Gonadotropin releasing hormone epitope (2-10
 FT aa)"
 FT Modified-site 47
 FT /note= "Amidated glycine or glycineamide"
 XX
 XX WO200185763-A2.
 XX
 PD 15-NOV-2001.
 XX
 XX 04-MAY-2001; 2001WO-US014363.
 PF

XX 05-MAY-2000; 2000US-0202328P.
 XX
 PA (APHT-) APHTON CORP.
 XX
 PI Grimes S, Michaeli D, Stevens VC;
 XX WPI; 2002-049440/06.
 XX
 XX Novel synthetic immunogen for inducing immune response against
 XX gonadotropin releasing hormone, comprises fusion peptide having
 XX promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 XX or its analog.
 XX
 PS Claim 11; Page 11; 43pp; English.
 XX
 CC The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX
 XX Sequence 47 AA;
 SQ
 Query Match 99.4%; Score 159; DB 5; Length 47;
 Best Local Similarity 100.0%; Pred. NO. 1.5e-16; Mismatches 0; Indels 0; Gaps 0;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KLLSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30
 |||||
 DB 17 KLLSEIKGVIVHRLEGVEGSPSLHWSYGLRP 46
 |||||
 RESULT 3
 AAR62705
 ID AAR62705 standard; peptide; 25 AA.
 XX
 AC AAR62705;
 XX
 DT 25-MAR-2003 (revised)
 DT 10-SEP-1995 (first entry)
 XX
 DE LHRH-containing immunogenic peptide.
 XX
 KW Helper T cell epitope; universal immune stimulator; invasin; hapten;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW measles virus F protein.
 XX
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT Domain 1..15
 FT /note= "measles virus F protein helper T cell epitope"
 FT Domain 16..25
 FT /note= "LHRH hapten"
 XX
 XX WO9425060-A1.
 XX
 PD 10-NOV-1994.
 XX
 XX 28-APR-1994; 94WO-US004832.
 XX
 XX 27-APR-1993; 93US-00057166.
 XX
 XX 14-APR-1994; 94US-0029275.
 XX

PA (LADD/) LADD A E.
 PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 XX
 PI Ladd AE, Wang CY, Zamb T;
 XX
 DR WPI; 1994-357910/44.
 XX
 XX Immunogenic luteinising hormone releasing hormone peptide(s) - that
 PT suppress LHRH activity in males and females.
 XX
 PS Claim 8; Page 84; 213pp; English.
 XX
 CC Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The stimulator
 CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
 CC immune response to the coupled peptide in members of a heterogeneous
 CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
 CC sequence from the invasive protein of Yersinia. Spacer amino acid
 CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
 CC domains and between the immune stimulator and hapten components. When the
 CC hapten is LHRH, then optionally the invasin domain can be omitted from
 CC the immune stimulator component. The present sequence represents an LHRH-
 CC containing, invasin-free immunogenic peptide as above which can be used
 CC as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-
 CC dependent carcinoma, prostatic carcinoma, testicular carcinoma,
 CC endometriosis, benign uterine tumours, recurrent functional ovarian
 CC cysts, (severe) premenstrual syndrome or oestrogen-dependent breast
 CC cancer, or for induction of infertility. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX
 SQ Sequence 25 AA;
 Query Match 73.1%; Score 117; DB 2; Length 25;
 Best Local Similarity 85.7%; Pred. NO. 1.7e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
 QY 3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
 DB 1 LSEIKGVIVHRLEGVE-----HWSYGLRP 24
 RESULT 4
 AAR62708
 ID AAR62708 standard; peptide; 42 AA.
 XX
 AC AAR62708;
 XX
 XX 25-MAR-2003 (revised)
 DT 10-SEP-1995 (first entry)
 DE LHRH-containing immunogenic peptide.
 XX
 KW Helper T cell epitope; universal immune stimulator; invasin; hapten;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW measles virus F protein.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Key 1..15
 FT Domain /note= "measles virus F protein helper T cell epitope"
 FT Domain 16..30
 FT Domain /note= "measles virus F protein helper T cell epitope"
 FT Domain 33..42
 FT Domain /note= "LHRH hapten"
 XX
 PN WO9425060-A1.
 XX
 PD 10-NOV-1994.

XX 28-APR-1994; 94WO-US004832.
 XX
 PR 27-APR-1993; 93US-00057166.
 PR 14-APR-1994; 94US-00229275.
 XX
 PA (LADD/) LADD A E.
 PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 XX
 PI Ladd AE, Wang CY, Zamb T;
 XX
 XX WPI; 1994-357910/44.
 XX
 XX Immunogenic luteinising hormone releasing hormone peptide(s) - that
 PT suppress LHRH activity in males and females.
 XX
 PS Claim 8; Page 86; 213pp; English.
 XX
 CC Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The stimulator
 CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
 CC immune response to the coupled peptide in members of a heterogeneous
 CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
 CC sequence from the invasive protein of Yersinia. Spacer amino acid
 CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
 CC domains and between the immune stimulator and hapten components. When the
 CC hapten is LHRH, then optionally the invasin domain can be omitted from
 CC the immune stimulator component. The present sequence represents an LHRH-
 CC containing, invasin-free immunogenic peptide as above which can be used
 CC as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-
 CC dependent carcinoma, prostatic carcinoma, testicular carcinoma,
 CC endometriosis, benign uterine tumours, recurrent functional ovarian
 CC cysts, (severe) premenstrual syndrome or oestrogen-dependent breast
 CC cancer, or for induction of infertility. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX
 SQ Sequence 42 AA;
 Query Match 73.1%; Score 117; DB 2; Length 42;
 Best Local Similarity 82.8%; Pred. NO. 3.2e-10;
 Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
 QY 2 LLEIKGVIVHRLEGVEGPSLHWSYGLRP 30
 DB 15 VLSEIKGVIVHRLEGVGGE--HWSYGLRP 41
 RESULT 5
 AAR62707
 ID AAR62707 standard; peptide; 27 AA.
 XX
 AC AAR62707;
 XX
 XX 25-MAR-2003 (revised)
 DT 10-SEP-1995 (first entry)
 DE LHRH-containing immunogenic peptide.
 XX
 KW Helper T cell epitope; universal immune stimulator; invasin; hapten;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW measles virus F protein.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Key 1..15
 FT Domain /note= "measles virus F protein helper T cell epitope"
 FT Domain 16..27
 FT Domain /note= "LHRH hapten"

XX PN W09425060-A1.
 XX PD 10-NOV-1994.
 XX PF 28-APR-1994; 94WO-US004832.
 XX PR 27-APR-1993; 93US-00057166.
 XX PR 14-APR-1994; 94US-00229275.
 XX PA (LADD/) LADD A B.
 XX PA (WANG/) WANG C Y.
 XX PA (ZAMB/) ZAMB T.
 XX PI Ladd AE, Wang CY, Zamb T;
 XX PX WPI; 1994-357910/44.
 XX DR Immunogenic luteinizing hormone releasing hormone peptide(s) - that suppress LHRH activity in males and females.
 XX PT Claim 8, 12; Page 86; 213pp; English.
 XX PS Synthetic immunogenic peptides are provided in which a universal immune stimulator is linked to a peptide or protein hapten containing B cell and/or cytotoxic T lymphocyte epitopes, giving a product which causes potent immune responses to the coupled peptide or protein. The stimulator consists of (A) a promiscuous helper T cell epitope (Th) which elicits an immune response to the coupled peptide in members of a heterogeneous population expressing diverse HLA phenotypes, and (B) an adjuvant peptide sequence from the invasive protein of Yersinia, Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the invasin and Th domains and between the immune stimulator and hapten components. When the hapten is LHRH, then optionally the invasin domain can be omitted from the immune stimulator component. The present sequence represents an LHRH-containing, invasin-free immunogenic peptide as above which can be used as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-dependent carcinoma, prostatic carcinoma, testicular carcinoma, endometriosis, benign uterine tumours, recurrent functional ovarian cysts, (severe) premenstrual syndrome or oestrogen-dependent breast cancer, or for induction of infertility. This sequence is particularly preferred. (Updated on 25-MAR-2003 to correct FN field.)
 XX SQ Sequence 27 AA;
 Query Match 72.5%; Score 116; DB 2; Length 27;
 Best Local Similarity 85.7%; Pred. No. 2.7e-10;
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
 QY 3 LSEIKGVIVHRLGVGSPSLHWSYGLRP 30
 DB 1 LSEIKGVIVHRLGVGSGE--HWSYGLRP 26
 RESULT 6
 AAY68567
 ID AAY68567 standard; peptide; 27 AA.
 XX AC AAY68567;
 XX DT 05-MAY-2000 (first entry)
 XX DE Peptide immunogen comprising a Th epitope and LHRH target antigen.
 XX KW Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH; luteinizing hormone-releasing hormone; spermatogenesis; ovulation;
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KW breast cancer; endometriosis; boar taint; meat quality; chimera; immunocastration.
 XX KW Measles virus.
 OS Unidentified.

OS Chimeric.
 XX Key Location/Qualifiers
 XX PH 1..15
 XX FT Peptide /note= "helper Th epitope AAY68540"
 XX FT Peptide 16..17
 XX FT Peptide /note= "spacer"
 XX FT Peptide 18..27
 XX FT Peptide /note= "LHRH antigenic epitope AAY68566"
 XX PN W09966952-A1.
 XX PX 29-DEC-1999.
 XX PF 21-JUN-1999; 99WO-US013960.
 XX PR 20-JUN-1998; 98US-00100414.
 XX PA (UNBI-) UNITED BIOMEDICAL INC.
 XX PX Wang CY;
 XX PX WPI; 2000-160562/14.
 XX DR New peptide immunogen containing luteinizing hormone-releasing hormone antigen site and helper T cell epitope, for e.g. contraception and treatment of cancer.
 XX PS Example 1; Page 63; 102pp; English.
 XX CC The present sequence represents a peptide immunogen comprising a helper T cell (Th) epitope of the F protein of the Measles virus and a target antigen, luteinizing hormone-releasing hormone (LHRH). The peptide immunogens cause induction of a specific immune response to LHRH which is involved in regulation of spermatogenesis, ovulation, oestrus, sexual development and secretion of sex hormones. Provision of a promiscuous T helper epitope (which is functional in genetically diverse subjects) provides optimum immunogenicity to the B cell epitopes of the target antigen and thus high antibody titres against the target antigen. The peptide immunogens of the invention are used to vaccinate against mammalian LHRH, for use as (reversible) contraceptive; control of hormone -dependent tumours (cancer of prostate or breast, also endometriosis); to prevent boar taint (and improve meat quality) and for immunocastration
 XX SQ Sequence 27 AA;
 Query Match 72.5%; Score 116; DB 3; Length 27;
 Best Local Similarity 85.7%; Pred. No. 2.7e-10;
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
 QY 3 LSEIKGVIVHRLGVGSPSLHWSYGLRP 30
 DB 1 LSEIKGVIVHRLGVGSGE--HWSYGLRP 26
 RESULT 7
 AAY91156
 ID AAY91156 standard; peptide; 27 AA.
 XX AC AAY91156;
 XX DT 12-SEP-2003 (revised)
 XX DT 22-MAY-2000 (first entry)
 XX DE MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:36.
 XX KW Promiscuous T-cell epitope; measles virus F protein; MVF; hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinizing hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
 KW cholesterol ester transport protein; anti-arteriosclerotic.

XX Measles virus.
 OS Rattus sp.
 OS Chimeric.
 XX WO9966957-A2.
 XX 29-DEC-1999.
 PD 21-JUN-1999; 99WO-US013975.
 XX 20-JUN-1998; 98US-00100412.
 XX (UNBI-) UNITED BIOMEDICAL INC.
 XX Wang CY;
 XX WPI; 2000-160564/14.
 XX
 XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus.
 XX
 PS Example 1; Page 77; 129pp; English.
 XX
 CC The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesterol ester transport
 CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone-
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and immunocastration
 CC ; for promoting the growth of animals; or for treating allergies or
 CC arteriosclerosis. Incorporation of a promiscuous Th (functional in
 CC genetically diverse subjects) into an immunogen improves capacity to
 CC induce a strong T helper cell-mediated immune response, resulting in
 CC production of antibodies against a target antigen. Th can replace carrier
 CC proteins and pathogen-derived T helper epitopes. Sequence AAY91121
 CC represents a promiscuous T helper epitope from the measles virus F (WVF)
 CC protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246
 CC represent synthetic Th epitopes based on the WVF Th epitope. Sequence
 CC AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)
 CC surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes
 CC derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-
 CC Y91244 are antigenic peptides comprising an LHRH sequence joined to a
 CC promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide
 CC used in these LHRH antigenic peptides. AAY91200 is somatostatin, and a Th
 CC AAY91201-Y91207 are antigenic peptides comprising somatostatin, and a Th
 CC epitope. Somatostatin immunogens may be used to promote growth in
 CC livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and
 CC AAY91209-Y90211 are HIV Th epitopes/CD4 CDR2 antigenic peptides which may
 CC be used to prevent HIV infection of T cells. AAY90212 is a modified
 CC version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219
 CC are Th epitope/IgE CH3 antigenic peptides which may be used in the
 CC treatment of allergies. AAY91220 is a peptide derived from foot and mouth
 CC disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this
 CC peptide and a Th epitope. AAY91223 is a Plasmodium falciparum
 CC circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS
 CC antigen and an WVF Th epitope and may be used in a malaria vaccine.
 CC AAY91228-Y91231 represent CETP-derived peptides and AAY91232-Y91241 are
 CC immunogens comprising a CETP peptide and a Th epitope which may be used
 CC to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247
 CC and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-
 CC Y91251 and AAY91258-Y91273 are antigenic peptides comprising WVF Th and
 CC HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1
 CC vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory
 CC invasive protein epitope from Yersinia species, and hinge spacer peptide,
 CC both of which may optionally be used in the antigenic peptides of the

CC invention. (Updated on 12-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 27 AA;
 Query Match 72.5%; Score 116; DB 3; Length 27;
 Best Local Similarity 85.7%; Pred. No. 2.7e-10;
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
 QY 3 LSEIKGVVHRLEGVGSPSLHWSYGLRP 30
 |||||
 DB 1 LSEIKGVVHRLEGVGGE--HWSYGLRP 26
 |||||
 RESULT 8
 AAR62721
 ID AAR62721 standard; peptide; 45 AA.
 XX
 AC AAR62721;
 XX
 DT 25-MAR-2003 (revised)
 DT 10-SEP-1995 (first entry)
 XX
 DE LHRH-containing immunogenic peptide.
 XX
 KW Helper T cell epitope; universal immune stimulator; invasin; haptens;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW measles virus F protein.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..16 /note= "invasin domain"
 FT Domain 19..33 /note= "measles virus F protein helper T cell epitope"
 FT Domain 36..45 /note= "LHRH haptens"
 FT
 XX WO9425060-A1.
 XX 10-NOV-1994.
 XX
 XX 28-APR-1994; 94WO-US004832.
 XX
 XX 27-APR-1993; 93US-00057166.
 XX 14-APR-1994; 94US-00229275.
 XX (LADD//) LADD A E.
 XX (WANG//) WANG C Y.
 XX (ZAMB//) ZAMB T.
 XX Ladd AE, Wang CY, Zamb T;
 XX WPI; 1994-357910/44.
 XX
 XX Immunogenic luteinising hormone releasing hormone peptide (s) - that
 XX suppress LHRH activity in males and females.
 PS Claim 8; Page 88; 213pp; English.
 XX
 CC Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein haptens containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The stimulator
 CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
 CC immune response to the coupled peptide in members of a heterogeneous
 CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
 CC sequence from the invasin protein of Yersinia. Spacer amino acid
 CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
 CC domains and between the immune stimulator and haptens components. When the
 CC haptens is LHRH, then optionally the invasin domain can be omitted from
 CC the immune stimulator component. The present sequence represents an LHRH-

CC containing immunogenic peptide as above which can be used as a potent
 CC vaccine for treating e.g. prostatic hyperplasia, androgen-dependent
 CC carcinoma, prostatic carcinoma, testicular carcinoma, endometriosis,
 CC benign uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX

SQ Sequence 45 AA;

Query Match 72.5%; Score 116; DB 2; Length 45;
 Best Local Similarity 85.7%; Pred. No. 5e-10;
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLEGVGPGLHWSYGLRP 30
 DB 19 LSEIKGVIVHRLEGVGPGLHWSYGLRP 44

RESULT 9
 ADD89949
 ID ADD89949 standard; protein; 45 AA.

XX
 AC ADD89949;

XX DT 29-JAN-2004 (first entry)

XX LHRH peptide used in immunostimulant complex for prostate cancer vaccine.

DE Immunostimulant; vaccine; human; immunogen; LHRH; immunotherapy;
 DE prostate cancer.

XX Synthetic.

OS Homo sapiens.

XX WO2003068169-A2.

XX 21-AUG-2003.

XX 14-FEB-2003; 2003WO-US004711.

XX 14-FEB-2002; 2002US-00076674.

PR 31-JAN-2003; 2003US-00076674.

XX (UNBI-) UNITED BIOMEDICAL INC.

PA Sokoll KK;

PI WPI; 2003-778890/73.

XX Stabilized immunostimulating complex, useful for vaccination, e.g.
 PT against human immune deficiency viruses, comprises cationic peptide
 PT immunogen and anionic oligonucleotide.

XX Claim 17; SEQ ID NO 9; 159pp; English.

XX The present sequence is that of a synthetic immunogenic peptide derived
 CC from human LHRH. This is an example of peptides that can be used in
 CC claimed immunostimulatory complexes of the invention that are
 CC specifically adapted to act as adjuvant and as peptide immunogen
 CC stabiliser. The complexes comprise a CpG oligonucleotide and a
 CC biologically active peptide immunogen. The complex is particulate and can
 CC efficiently present peptide immunogens to the cells of the immune system
 CC to produce an immune response. The complexes may be prepared with various
 CC ratios of peptides to CpG oligonucleotides to provide different physical
 CC properties, such as the size of the microparticle. An immunostimulatory
 CC complex comprising the present LHRH derived peptide can be used in a
 CC vaccine for prostate cancer.

XX Sequence 45 AA;

Query Match 72.5%; Score 116; DB 7; Length 45;
 Best Local Similarity 85.7%; Pred. No. 5e-10;
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLEGVGPGLHWSYGLRP 30
 DB 19 LSEIKGVIVHRLEGVGPGLHWSYGLRP 44

RESULT 10

AAV91163

ID AAV91163 standard; peptide; 27 AA.

XX AAV91163;

XX 12-SEP-2003 (revised)

DT 22-MAY-2000 (first entry)

DE Modified MWF Th epitope/LHRH antigenic peptide, SEQ ID NO:43.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;

KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;

KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;

KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;

KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;

KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;

KW cholesterol ester transport protein; anti-arteriosclerotic.

XX Measles virus.

OS Rattus sp.

OS Chimeric.

XX WO9966957-A2.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US013975.

XX 20-JUN-1998; 98US-00100412.

PA (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY;

XX WPI; 2000-160564/14.

XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus.

XX Example 1; Page 80; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesterol ester transport
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone-
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and immunocastration
 CC ; for promoting the growth of animals; or for treating allergies or
 CC arteriosclerosis. Incorporation of a promiscuous Th (functional in
 CC genetically diverse subjects) into an immunogen improves capacity to
 CC induce a strong T helper cell-mediated immune response, resulting in
 CC production of antibodies against a target antigen. Th can replace carrier
 CC proteins and pathogen-derived T helper epitopes. Sequence AAY91163
 CC represents a promiscuous T helper epitope from the measles virus F (MVF)
 CC protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246
 CC represent synthetic Th epitopes based on the MWF Th epitope. Sequence
 CC AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)
 CC surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes
 CC derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-

XX 12-SEP-2003 (revised)
 DT 22-MAY-2000 (first entry)
 XX Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:41.
 DE Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;
 KW cholesterol ester transport protein; anti-arteriosclerotic.
 XX Measles virus.
 OS Rattus sp.
 OS Chimeric.
 XX WO9966957-A2.
 PN 29-DEC-1999.
 XX 21-JUN-1999; 99WO-US013975.
 XX 20-JUN-1998; 98US-00100412.
 XX (UNBI-) UNITED BIOMEDICAL INC.
 XX Wang CY;
 XX WPI; 2000-160564/14.
 XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus.
 XX Example 1; Page 79; 129pp; English.
 PS The invention relates to novel promiscuous T helper cell epitopes (Th),
 XX and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesterol ester transport
 CC protein (CPTP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and immunocastration)
 CC ; for promoting the growth of animals; or for treating allergies or
 CC arteriosclerosis. Incorporation of a promiscuous Th (functional in
 CC genetically diverse subjects) into an immunogen improves capacity to
 CC induce a strong T helper cell-mediated immune response, resulting in
 CC production of antibodies against a target antigen. Th can replace carrier
 CC proteins and pathogen-derived T helper epitopes. Sequence AAY91121
 CC represents a promiscuous T helper epitope from the measles virus F (MVF)
 CC protein and sequences AAY91121-Y91142, AAY91226 and AAY91245-Y91246
 CC represent synthetic Th epitopes based on the MVF Th epitope. Sequence
 CC AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)
 CC surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes
 CC derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-
 CC Y91244 are antigenic peptides comprising an LHRH sequence joined to a
 CC promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide
 CC used in these LHRH antigenic peptides. AAY91200 is somatostatin, and
 CC AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th
 CC epitope. Somatostatin immunogens may be used to promote growth in
 CC livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and
 CC AAY91209-Y90211 are MVF Th epitope/CD4 CDR2 antigenic peptides which may
 CC be used to prevent HIV infection of T cells. AAY90212 is a modified
 CC version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219
 CC are Th epitope/IgE CH3 immunogenic peptides which may be used in the
 CC treatment of allergies. AAY91220 is a peptide derived from foot and mouth

CC disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this
 CC peptide and a Th epitope. AAY91223 is a Plasmodium falciparum
 CC circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS
 CC antigen and an MVF Th epitope and may be used in a malaria vaccine.
 CC AAY91228-Y91231 represent CPTP-derived peptides and AAY91232-Y91241 are
 CC immunogens comprising a CPTP peptide and a Th epitope which may be used
 CC to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247
 CC and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-
 CC Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVF Th and
 CC HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1
 CC vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory
 CC invasion protein epitope from *Yersinia* species, and hinge spacer peptide,
 CC both of which may optionally be used in the antigenic peptides of the
 CC invention. (Updated on 12-SEP-2003 to standardise OS field)
 XX Sequence 27 AA;
 SQ
 Query Match 67.5%; Score 108; DB 3; Length 27;
 Best Local Similarity 71.4%; Pred. No. 4.4e-09;
 Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;
 QY 3 LSEIKGVIVHRLGVEGPESLHWSYGLRP 30
 DB 1 ISEIKGVIVHRLGVEGPE--HWSYGLRP 26
 RESULT 13
 AAY91167
 ID AAY91167 standard; peptide; 27 AA.
 XX
 AC AAY91167;
 XX
 DT 12-SEP-2003 (revised)
 DT 22-MAY-2000 (first entry)
 XX Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:47.
 XX Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;
 KW cholesterol ester transport protein; anti-arteriosclerotic.
 XX Measles virus.
 OS Rattus sp.
 OS Chimeric.
 XX WO9966957-A2.
 PN 29-DEC-1999.
 XX 21-JUN-1999; 99WO-US013975.
 XX 20-JUN-1998; 98US-00100412.
 XX (UNBI-) UNITED BIOMEDICAL INC.
 XX Wang CY;
 XX WPI; 2000-160564/14.
 XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus.
 XX Example 1; Page 81; 129pp; English.
 PS The invention relates to novel promiscuous T helper cell epitopes (Th),
 XX and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesterol ester transport
 CC protein (CPTP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and immunocastration)
 CC ; for promoting the growth of animals; or for treating allergies or
 CC arteriosclerosis. Incorporation of a promiscuous Th (functional in
 CC genetically diverse subjects) into an immunogen improves capacity to
 CC induce a strong T helper cell-mediated immune response, resulting in
 CC production of antibodies against a target antigen. Th can replace carrier
 CC proteins and pathogen-derived T helper epitopes. Sequence AAY91121
 CC represents a promiscuous T helper epitope from the measles virus F (MVF)
 CC protein and sequences AAY91121-Y91142, AAY91226 and AAY91245-Y91246
 CC represent synthetic Th epitopes based on the MVF Th epitope. Sequence
 CC AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)
 CC surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes
 CC derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-
 CC Y91244 are antigenic peptides comprising an LHRH sequence joined to a
 CC promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide
 CC used in these LHRH antigenic peptides. AAY91200 is somatostatin, and
 CC AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th
 CC epitope. Somatostatin immunogens may be used to promote growth in
 CC livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and
 CC AAY91209-Y90211 are MVF Th epitope/CD4 CDR2 antigenic peptides which may
 CC be used to prevent HIV infection of T cells. AAY90212 is a modified
 CC version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219
 CC are Th epitope/IgE CH3 immunogenic peptides which may be used in the
 CC treatment of allergies. AAY91220 is a peptide derived from foot and mouth

CC peptide and a Th epitope. AA9191223 is a Plasmodium falciparum
CC circumsporozoite (CS) target antigen, and AA9191224-9191225 comprise the CS
CC antigen and an MZF Th epitope and may be used in a malaria vaccine.
CC AA9191228-9191231 represent CERP-derived peptides and AA9191232-9191241 are
CC immunogens comprising a CERP peptide and a Th epitope which may be used
CC to prevent or treat arteriosclerosis and cardiovascular disease. AA9191247
CC and AA9191252-9191257 are HIV-1 neutralising B-cell epitopes, and AA9191248-
CC 9191251 and AA9191258-9191273 are antigenic peptides comprising MZF Th and
CC HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1
CC vaccine. AA9191198 and AA9191199 are respectively an immunostimulatory
CC invasion protein epitope from Yersinia species, and hinge spacer peptide,
CC both of which may optionally be used in the antigenic peptides of the
CC invention. (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 45 AA; 67.5%; Score 108; DB 3; Length 45;
Query Match
Best Local Similarity 71.4%; Pred. No. 8.2e-09;
Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1

QY 3 LSEIKGVIVHRLLEGVEGFSLEWSYGLRP 30
:|||||:|||||:
Db 19 ISEIKGVIVHRIETGIGGE--HWSYGLRP 44

Search completed: March 10, 2004, 09:12:09
Job time : 47.6809 secs

The invention relates to novel promiscuous T helper cell epitopes (Th), and immunogenic peptides comprising the Th epitopes of the invention along with B cell epitopes. The Th epitopes and peptide immunogens containing them, are used to induce a T helper cell response, specifically against Plasmodium falciparum, cholesterol ester transport protein (CEPT) or HIV epitopes, but more generally against any pathogen, immunoreactive self-antigen or tumour antigen. The Th epitopes and peptide immunogens may be used for prevention and/or treatment of infections (HIV, foot-and-mouth disease or malaria); for cancer immunotherapy; for inhibition of the action of luteinizing hormone releasing hormone (LHRH) for contraception, treatment of hormone-dependent cancer, prevention of boar taint in meat, and immunocastration; for promoting the growth of animals, or for treating allergies or arteriosclerosis. Incorporation of a promiscuous Th (functional in genetically diverse subjects) into an immunogen improves capacity to induce a strong T helper cell-mediated immune response, resulting in production of antibodies against a target antigen. Th can replace carrier proteins and pathogen-derived T helper epitopes. Sequence AA91121 represents a promiscuous T helper epitope from the measles virus F (MVP) protein and sequences AA91122-Y91142, AA91226 and AA91245-Y91246 represent synthetic Th epitopes based on the MVP Th epitope. Sequence AA91143 represents a promiscuous Th epitope from hepatitis B virus (HBV) surface antigen, and sequences AA91144-Y91155 are synthetic epitopes derived from this HBV epitope. AA91156-Y91196, AA91227 and AA91242-Y91244 are antigenic peptides comprising an LHRH sequence joined to a promiscuous Th epitope. AA91219 is the LHRH target antigenic peptide used in these LHRH antigenic peptides. AA91200 is somatostatin, and AA91201-Y91207 are antigenic peptides comprising somatostatin and a Th epitope. Somatostatin immunogens may be used to promote growth in livestock. AA91208 is a human CD4 CD82-like domain antigenic site, and AA91209-Y90211 are MVM Th epitope/CD4 CD82 antigenic peptides which may be used to prevent HIV infection of T cells. AA912012 is a modified version of a human IGE (immunoglobulin E) CH3 domain, and AA91213-Y90219 are Th epitope/IGE CH3 antigenic peptides which may be used in the treatment of allergies. AA91220 is a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid protein and AA91221-Y91222 comprise this

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 4.71595 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-6

Perfect score: 29
Sequence: 1 SSGPSSL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|---------------------|
| 1 | 29 | 100.0 | 6 | 9 | US-09-848-834A-6 |
| 2 | 29 | 100.0 | 8 | 9 | US-09-848-834A-7 |
| 3 | 29 | 100.0 | 31 | 9 | US-09-848-834A-15 |
| 4 | 29 | 100.0 | 34 | 9 | US-09-848-834A-13 |
| 5 | 29 | 100.0 | 36 | 9 | US-09-848-834A-16 |
| 6 | 29 | 100.0 | 37 | 9 | US-09-848-834A-14 |
| 7 | 29 | 100.0 | 46 | 9 | US-09-848-834A-19 |
| 8 | 29 | 100.0 | 47 | 9 | US-09-848-834A-17 |
| 9 | 29 | 100.0 | 50 | 9 | US-09-848-834A-18 |
| 10 | 29 | 100.0 | 51 | 9 | US-09-848-834A-20 |
| 11 | 29 | 100.0 | 53 | 9 | US-09-864-761-47649 |
| 12 | 29 | 100.0 | 75 | 9 | US-09-731-872-425 |
| 13 | 29 | 100.0 | 75 | 10 | US-09-876-997-425 |
| 14 | 29 | 100.0 | 200 | 9 | US-09-870-756-38 |
| 15 | 29 | 100.0 | 200 | 9 | US-09-874-585B-38 |

| | | | | | | |
|----|----|-------|------|----|--------------------|--------------------|
| 16 | 29 | 100.0 | 492 | 14 | US-10-156-761-9393 | Sequence 9393, Ap |
| 17 | 29 | 100.0 | 1197 | 9 | US-09-738-626-6493 | Sequence 6493, Ap |
| 18 | 29 | 100.0 | 3186 | 15 | US-10-210-130-34 | Sequence 34, Appl |
| 19 | 29 | 100.0 | 3208 | 15 | US-10-210-130-38 | Sequence 38, Appl |
| 20 | 29 | 100.0 | 3252 | 15 | US-10-210-130-36 | Sequence 36, Appl |
| 21 | 29 | 100.0 | 3262 | 15 | US-10-379-381-4 | Sequence 4, Appl |
| 22 | 29 | 100.0 | 3268 | 15 | US-10-379-381-2 | Sequence 2, Appl |
| 23 | 27 | 93.1 | 122 | 10 | US-09-764-891-3273 | Sequence 3273, Ap |
| 24 | 27 | 93.1 | 606 | 14 | US-10-225-567A-438 | Sequence 438, Appl |
| 25 | 27 | 93.1 | 645 | 9 | US-09-796-338A-2 | Sequence 2, Appl |
| 26 | 27 | 93.1 | 645 | 10 | US-09-977-418-28 | Sequence 28, Appl |
| 27 | 27 | 93.1 | 645 | 10 | US-09-977-418-46 | Sequence 46, Appl |
| 28 | 27 | 93.1 | 645 | 10 | US-09-977-033A-28 | Sequence 28, Appl |
| 29 | 27 | 93.1 | 645 | 10 | US-09-977-033A-46 | Sequence 46, Appl |
| 30 | 27 | 93.1 | 645 | 10 | US-09-977-751C-28 | Sequence 28, Appl |
| 31 | 27 | 93.1 | 645 | 10 | US-09-977-751C-46 | Sequence 46, Appl |
| 32 | 27 | 93.1 | 645 | 10 | US-09-977-639A-28 | Sequence 28, Appl |
| 33 | 27 | 93.1 | 645 | 10 | US-09-977-639A-46 | Sequence 46, Appl |
| 34 | 27 | 93.1 | 645 | 11 | US-09-977-819B-28 | Sequence 28, Appl |
| 35 | 27 | 93.1 | 645 | 11 | US-09-977-819B-46 | Sequence 46, Appl |
| 36 | 27 | 93.1 | 645 | 14 | US-10-282-837-2 | Sequence 2, Appl |
| 37 | 27 | 93.1 | 645 | 14 | US-10-145-586-2 | Sequence 2, Appl |
| 38 | 27 | 93.1 | 683 | 10 | US-09-977-418-26 | Sequence 26, Appl |
| 39 | 27 | 93.1 | 683 | 10 | US-09-977-033A-26 | Sequence 26, Appl |
| 40 | 27 | 93.1 | 683 | 10 | US-09-977-751C-26 | Sequence 26, Appl |
| 41 | 27 | 93.1 | 683 | 10 | US-09-977-639A-26 | Sequence 26, Appl |
| 42 | 27 | 93.1 | 683 | 11 | US-09-977-819B-26 | Sequence 26, Appl |
| 43 | 27 | 93.1 | 890 | 9 | US-09-828-366-7 | Sequence 7, Appl |
| 44 | 27 | 93.1 | 890 | 9 | US-09-909-320-49 | Sequence 49, Appl |
| 45 | 27 | 93.1 | 890 | 9 | US-09-909-088B-49 | Sequence 49, Appl |

ALIGNMENTS

RESULT 1
US-09-848-834A-6
; Sequence 6, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCES: 1102865-0047 US/09/848,834A
; CURRENT APPLICATION NUMBER: 2001-05-04
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-09-848-834A-6

Query Match 100.0%; Score 29; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSSL 6
Db 1 SSGPSSL 6

RESULT 2
US-09-848-834A-7
; Sequence 7, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens

FILE REFERENCE: 1102865-0047
CURRENT APPLICATION NUMBER: US/09/848,834A
PRIOR FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: 60/202,328
PRIOR FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patentin version 3.0
SEQ ID NO 7
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic peptide
US-09-848-834A-7

Query Match 100.0%; Score 29; DB 9; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.1e-05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
DB 1 SSGPSL 6

RESULT 3
US-09-848-834A-15
Sequence 15, Application US/09848834A
Patent No. US20020076416A1
GENERAL INFORMATION:
APPLICANT: Aphton Corporation
TITLE OF INVENTION: Chimeric Peptide Immunogens
FILE REFERENCE: 1102865-0047
CURRENT APPLICATION NUMBER: US/09/848,834A
CURRENT FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: 60/202,328
PRIOR FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patentin version 3.0
SEQ ID NO 15
LENGTH: 31
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gn
OTHER INFORMATION: RH hormone linked by a spacer to amino sequence 830-844 of the Te
OTHER INFORMATION: tanus toxoid precursor (Tentoxylisin)
NAME/KEY: MOD RES
LOCATION: (1)..(1)
OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
NAME/KEY: PEPTIDE
LOCATION: (1)..(10)
OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
NAME/KEY: PEPTIDE
LOCATION: (11)..(16)
OTHER INFORMATION: Spacer peptide
NAME/KEY: PEPTIDE
LOCATION: (17)..(31)
OTHER INFORMATION: (Tentoxylisin)
US-09-848-834A-15

Query Match 100.0%; Score 29; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
DB 11 SSGPSL 16

RESULT 4
US-09-848-834A-13
Sequence 13, Application US/09848834A

PATENT NO. US20020076416A1
GENERAL INFORMATION:
APPLICANT: Aphton Corporation
TITLE OF INVENTION: Chimeric Peptide Immunogens
FILE REFERENCE: 1102865-0047
CURRENT APPLICATION NUMBER: US/09/848,834A
CURRENT FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: 60/202,328
PRIOR FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patentin version 3.0
SEQ ID NO 13
LENGTH: 34
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the G
OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of
OTHER INFORMATION: he Measles virus fusion protein,
NAME/KEY: PEPTIDE
LOCATION: (1)..(10)
OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
NAME/KEY: PEPTIDE
LOCATION: (11)..(18)
OTHER INFORMATION: Spacer peptide
NAME/KEY: PEPTIDE
LOCATION: (19)..(34)
OTHER INFORMATION: Amino acid sequence 288-302 of the Measles
OTHER INFORMATION: virus fusion protein, F
NAME/KEY: MOD RES
LOCATION: (1)..(1)
OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
US-09-848-834A-13

Query Match 100.0%; Score 29; DB 9; Length 34;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
DB 11 SSGPSL 16

RESULT 5
US-09-848-834A-16
Sequence 16, Application US/09848834A
Patent No. US20020076416A1
GENERAL INFORMATION:
APPLICANT: Aphton Corporation
TITLE OF INVENTION: Chimeric Peptide Immunogens
FILE REFERENCE: 1102865-0047
CURRENT APPLICATION NUMBER: US/09/848,834A
CURRENT FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: 60/202,328
PRIOR FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patentin version 3.0
SEQ ID NO 16
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the C
OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 378-398 of
OTHER INFORMATION: the Plasmodium falciparum circumsporozoite (CSP) protein
NAME/KEY: MOD RES
LOCATION: (1)..(1)
OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
NAME/KEY: PEPTIDE
LOCATION: (1)..(10)
OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
NAME/KEY: PEPTIDE
LOCATION: (11)..(16)

OTHER INFORMATION: Spacer peptide
NAME/KEY: PEPTIDE
LOCATION: (17)..(36)
OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria
OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite
OTHER INFORMATION: (CSP) protein
US-09-848-834A-16

Query Match 100.0%; Score 29; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
DB 11 SSGPSL 16

RESULT 6

US-09-848-834A-14
Sequence 14, Application US/09848834A
Patent No. US20020076416A1

GENERAL INFORMATION:
APPLICANT: Apton Corporation
TITLE OF INVENTION: Chimeric Peptide Immunogens
FILE REFERENCE: 1102865-0047
CURRENT APPLICATION NUMBER: US/09/848,834A
CURRENT FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: 60/202,328
PRIOR FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.0
SEQ ID NO 14
LENGTH: 37
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the

OTHER INFORMATION: nRH hormone linked by a spacer to amino acid sequence 947-967 of
OTHER INFORMATION: the Tetanus toxoid precursor (Tentoxylisin)

NAME/KEY: MOD_RES
LOCATION: (1)..(1)
OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

NAME/KEY: PEPTIDE
LOCATION: (1)..(10)
OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone

NAME/KEY: PEPTIDE
LOCATION: (11)..(16)
OTHER INFORMATION: Spacer peptide
NAME/KEY: PEPTIDE
LOCATION: (17)..(37)

OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor
OTHER INFORMATION: (Tentoxylisin)

US-09-848-834A-14

Query Match 100.0%; Score 29; DB 9; Length 37;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
DB 11 SSGPSL 16

RESULT 7

US-09-848-834A-19

Sequence 19, Application US/09848834A
Patent No. US20020076416A1

GENERAL INFORMATION:

APPLICANT: Apton Corporation
TITLE OF INVENTION: Chimeric Peptide Immunogens
FILE REFERENCE: 1102865-0047
CURRENT APPLICATION NUMBER: US/09/848,834A
CURRENT FILING DATE: 2001-05-04

PRIOR APPLICATION NUMBER: 60/202,328
PRIOR FILING DATE: 2000-05-05

NUMBER OF SEQ ID NOS: 20

SOFTWARE: PatentIn version 3.0

SEQ ID NO 19

LENGTH: 46

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of huma
OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tetan
OTHER INFORMATION: toxoid precursor (Tentoxylisin) linked by a spacer to amino aci
OTHER INFORMATION: sequence 1-10 of GnRH

NAME/KEY: MOD_RES

LOCATION: (1)..(1)

OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

NAME/KEY: MOD_RES

LOCATION: (46)..(46)

OTHER INFORMATION: Amidated glycine or glycylamide

NAME/KEY: PEPTIDE

LOCATION: (1)..(10)

OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone

NAME/KEY: PEPTIDE

LOCATION: (11)..(16)

OTHER INFORMATION: Spacer peptide

NAME/KEY: PEPTIDE

LOCATION: (17)..(31)

OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
OTHER INFORMATION: (Tentoxylisin)

NAME/KEY: PEPTIDE

LOCATION: (32)..(37)

OTHER INFORMATION: Spacer peptide

NAME/KEY: PEPTIDE

LOCATION: (38)..(46)

OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone

US-09-848-834A-19

Query Match 100.0%; Score 29; DB 9; Length 46;

Best Local Similarity 100.0%; Pred. No. 11e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6

DB 11 SSGPSL 16

RESULT 8

US-09-848-834A-17

Sequence 17, Application US/09848834A

Patent No. US20020076416A1

GENERAL INFORMATION:

APPLICANT: Apton Corporation

TITLE OF INVENTION: Chimeric Peptide Immunogens

FILE REFERENCE: 1102865-0047

CURRENT APPLICATION NUMBER: US/09/848,834A

CURRENT FILING DATE: 2001-05-04

PRIOR APPLICATION NUMBER: 60/202,328

PRIOR FILING DATE: 2000-05-05

NUMBER OF SEQ ID NOS: 20

SOFTWARE: PatentIn version 3.0

SEQ ID NO 17

LENGTH: 47

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of

OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid se
OTHER INFORMATION: uence 2-10 of the GnRH hormone

NAME/KEY: MOD_RES

LOCATION: (1)..(1)

OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

NAME/KEY: MOD_RES

LOCATION: (47)..(47)
; OTHER INFORMATION: Amidated-glycine or glycineamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(18)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (19)..(34)
; OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein,
; NAME/KEY: PEPTIDE
; LOCATION: (35)..(38)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (39)..(47)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-17

Query Match 100.0%; Score 29; DB 9; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
Db 11 SSGPSL 16

RESULT 9
US-09-848-834A-18
; Sequence 18, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the Tet
; OTHER INFORMATION: anus toxoid precursor (Tentoxylisin) protein linked by a spacer
; OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD RES
; LOCATION: (50)..(50)
; OTHER INFORMATION: Amidated glycine or glycineamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(37)
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Tent
; OTHER INFORMATION: oxylysin
; NAME/KEY: PEPTIDE
; LOCATION: (38)..(41)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (42)..(50)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone

US-09-848-834A-18

Query Match 100.0%; Score 29; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
Db 11 SSGPSL 16

RESULT 10
US-09-848-834A-20
; Sequence 20, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of Plasmo
; OTHER INFORMATION: ium falciparum circumsporozoite (CSP) protein
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD RES
; LOCATION: (51)..(51)
; OTHER INFORMATION: Amidated glycine or glycineamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(36)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum
; OTHER INFORMATION: circumsporozoite (CSP) protein
; NAME/KEY: PEPTIDE
; LOCATION: (37)..(42)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (43)..(51)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-20

Query Match 100.0%; Score 29; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
Db 11 SSGPSL 16

RESULT 11
US-09-864-761-47649
; Sequence 47649, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aecmca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 47649
LENGTH: 53
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO Z83826.12
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.73
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 0.96
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.57
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.62
OTHER INFORMATION: SWISSPROT HIT: O52399, EVALUE 5.90e+00
OTHER INFORMATION: EST_HUMAN HIT: AA813575.1, EVALUE 3.00e-22
US-09-864-761-47649

Query Match 100.0%; Score 29; DB 9; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
|||||
DB 5 SSGPSL 10

RESULT 12
US-09-731-872-425
Sequence 425, Application US/09731872
Patent No. US20020102604A1
GENERAL INFORMATION:
APPLICANT: VIITANEN, PAUL VEIKKO
BACOT, KAREN ONLEY
JORDAN, DOUGLAS BRIAN
TITLE OF INVENTION: LUMAZINE SYNTHASE AND
RIBOFLAVIN SYNTHASE

APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Bougueleret, Lydie
APPLICANT: Jobert, Severin
TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78.US3.REG
CURRENT APPLICATION NUMBER: US/09/731,872
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 425
LENGTH: 75
TYPE: PRT
ORGANISM: Homo sapiens
US-09-731-872-425

Query Match 100.0%; Score 29; DB 9; Length 75;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
|||||
DB 41 SSGPSL 46

RESULT 13
US-09-876-997-425
Sequence 425, Application US/09876997
Publication No. US20030152921A1
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Bougueleret, Lydie
APPLICANT: Jobert, Severin
TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78.US4.CIP
CURRENT APPLICATION NUMBER: US/09/876,997
CURRENT FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: US 09/731,872
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 425
LENGTH: 75
TYPE: PRT
ORGANISM: Homo sapiens
US-09-876-997-425

Query Match 100.0%; Score 29; DB 10; Length 75;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
|||||
DB 41 SSGPSL 46

RESULT 14
US-09-870-756-38
Sequence 38, Application US/09870756
Patent No. US20020052023A1
GENERAL INFORMATION:
APPLICANT: VIITANEN, PAUL VEIKKO
BACOT, KAREN ONLEY
JORDAN, DOUGLAS BRIAN
TITLE OF INVENTION: LUMAZINE SYNTHASE AND
RIBOFLAVIN SYNTHASE

QY 1 SSGPSL 6
Db 79 SSGPSL 84
Search completed: March 10, 2004, 10:25:47
Job time : 4.71595 secs

NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESS: B. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA: US/09/870,756
APPLICATION NUMBER: US/09/870,756
FILING DATE: 31-May-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CL-1083
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-8112
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 200 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: M. grisea LS
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-870-756-38

Query Match 100.0%; Score 29; DB 9; Length 200;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
Db 79 SSGPSL 84

RESULT 15
US-09-874-585B-38
Sequence 38, Application US/09874585B
Patent No. US20020127670A1
GENERAL INFORMATION:
APPLICANT: Viltanen, Paul Veikko
APPLICANT: Jordan, Douglas Brain
APPLICANT: Bacot, Karen Onley
TITLE OF INVENTION: Riboflavin Synthase Genes and Enzymes and Methods of Use
FILE REFERENCE: CL1083 US DIV2
CURRENT APPLICATION NUMBER: US/09/874,585B
CURRENT FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: 09/874,585
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 08/912,218
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Microsoft Office 97
SEQ ID NO 38
LENGTH: 200
TYPE: PRT
ORGANISM: Magnaporthe grisea
US-09-874-585B-38

Query Match 100.0%; Score 29; DB 9; Length 200;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC spacer peptide used in the immunogen of the invention
 XX
 SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 5; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
 |||||
 DB 1 SSGPSL 6

RESULT 2
 AAU11418
 ID AAU11418 standard; peptide; 8 AA.
 XX
 AC AAU11418;
 XX
 DT 12-MAR-2002 (first entry)
 DE
 DE Synthetic spacer peptide #3.

XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 XX luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer;
 KW spacer peptide.

XX
 OS Synthetic.
 OS
 PN WO200185763-A2.
 XX
 PD 15-NOV-2001.

XX 04-MAY-2001; 2001WO-US014363.
 XX 05-MAY-2000; 2000US-0202328P.
 DR (APHT-) APHTON CORP.

XX Grimes S, Michaeli D, Stevens VC;
 XX WPI; 2002-049440/06.
 XX Novel synthetic immunogen for inducing immune response against
 FT Gonadotropin releasing hormone, comprises fusion peptide having
 FT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 FT or its analog.

XX Claim 10; Page 6; 43pp; English.

XX The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC spacer peptide used in the immunogen of the invention

XX Sequence 8 AA;

Query Match 100.0%; Score 29; DB 5; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSGPSL 6
 |||||
 DB 1 SSGPSL 6

RESULT 3
 AAU11426
 ID AAU11426 standard; peptide; 31 AA.

XX AC AAU11426;
 XX
 DT 12-MAR-2002 (first entry)
 DE
 DE Synthetic immunogen peptide 7.

XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX Clostridium tetani.
 OS Mammalia.
 OS Synthetic.
 OS Chimeric.

XX Key Location/Qualifiers
 FT Peptide 1..10
 FT Misc-difference 1 /note= "Gonadotropin releasing hormone epitope"

FT /label= OTHER
 FT Peptide 11..16 /note= "Pyro-glutamic acid or 5-oxo proline"

FT Peptide 17..31 /note= "Spacer peptide"

FT /note= "Tetanus toxoid sequence (830-844 aa)"
 XX WO200185763-A2.
 XX 15-NOV-2001.

XX 04-MAY-2001; 2001WO-US014363.
 XX 05-MAY-2000; 2000US-0202328P.
 XX (APHT-) APHTON CORP.

XX Grimes S, Michaeli D, Stevens VC;
 XX WPI; 2002-049440/06.
 XX Novel synthetic immunogen for inducing immune response against
 FT Gonadotropin releasing hormone, comprises fusion peptide having
 FT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 FT or its analog.

XX Claim 11; Page 10; 43pp; English.
 XX The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention

SQ Sequence 31 AA;
Query Match 100.0%; Score 29; DB 5; Length 31;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
| | | | |
Db 11 SSGPSL 16

RESULT 4
AAU11424
ID AAU11424 standard; peptide; 34 AA.
XX
AC AAU11424;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 5.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS Measles virus.
OS Mammalia.
OS Synthetic.
OS Chimeric.

XX Key Location/Qualifiers
FH Key 1..10
FT Peptide /note= "Gonadotrophin releasing hormone epitope"
FT Misc-difference 1
FT Peptide /label= OTHER
FT Peptide /note= "Other- Pyro- glutamic acid or 5-oxo proline"
FT Peptide /note= "Spacer peptide"
FT Peptide /note= "Measles virus fusion protein F epitope"
XX WO200185763-A2.
XX
PD 15-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014363.
XX
PR 05-MAY-2000; 2000US-0202328P.
XX
PA (APHT-) APHTON CORP.
XX
PI Grimes S, Michaeli D, Stevens VC;
XX
DR WPI; 2002-049440/06.
XX
PS Novel synthetic immunogen for inducing immune response against
PT gonadotrophin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
or its analog.
XX
PS Claim 11; Page 9; 43pp; English.

XX The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotrophin releasing hormone (GnRH) also known as
CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC useful inducing an immune response against GnRH in an animal subject, and
CC as such is useful as a contraceptive and in the treatment of diseases
CC such as cancer (of the breast, uterus and other gynaecological cancer),
CC endometriosis, uterine fibroids, benign prostatic hypertrophy and

CC prostate cancer. The immunogen is effective in eliciting high and
CC specific anti-GnRH antibody titres. The present sequence is a synthetic
CC immunogen of the invention
XX
SQ Sequence 34 AA;
Query Match 100.0%; Score 29; DB 5; Length 34;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
| | | | |
Db 11 SSGPSL 16

RESULT 5
AAU11427
ID AAU11427 standard; peptide; 36 AA.
XX
AC AAU11427;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 8.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS Plasmodium falciparum.
OS Mammalia.
OS Synthetic.
OS Chimeric.

XX Key Location/Qualifiers
FH Key 1..10
FT Peptide /note= "Gonadotrophin releasing hormone epitope"
FT Misc-difference 1
FT Peptide /label= OTHER
FT Peptide /note= "Other- Pyro- glutamic acid or 5-oxo proline"
FT Peptide /note= "Spacer peptide"
FT Peptide /note= "Malaria CSP protein (378-398 aa)"
XX WO200185763-A2.
XX
PD 15-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014363.
XX
PR 05-MAY-2000; 2000US-0202328P.
XX
PA (APHT-) APHTON CORP.
XX
PI Grimes S, Michaeli D, Stevens VC;
XX
DR WPI; 2002-049440/06.
XX
PS Novel synthetic immunogen for inducing immune response against
PT gonadotrophin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
or its analog.
XX
PS Claim 11; Page 10; 43pp; English.

XX The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotrophin releasing hormone (GnRH) also known as
CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic immunogen is

CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX
 SQ Sequence 36 AA;

Query Match 100.0%; Score 29; DB 5; Length 36;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
 Db 11 SSGPSL 16

RESULT 6
 AAU11425
 ID AAU11425 standard; peptide; 37 AA.

XX AC AAU11425;

XX DT 12-MAR-2002 (first entry)

XX DE Synthetic immunogen peptide 6.

XX Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX Clostridium tetani.

OS Mammalia.

OS Synthetic.

OS Chimeric.

XX Key Location/Qualifiers

FT Peptide 1..10

FT /note= "Gonadotrophin releasing hormone epitope"

FT Misc-difference 1

FT /label= OTHER

FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"

FT Peptide 11..16

FT /note= "Spacer peptide"

FT Peptide 17..37

FT /note= "Tetanus toxoid sequence (947-967 aa)"

FT WO200185763-A2.

PN 15-NOV-2001.

XX 04-MAY-2001; 2001WO-US014363.

XX 05-MAY-2000; 2000US-0202328P.

XX (APHT-) APHTON CORP.

XX Grimes S, Michaeli D, Stevens VC;

XX WPI; 2002-049440/06.

XX Novel synthetic immunogen for inducing immune response against
 PT Gonadotrophin releasing hormone, comprises fusion peptide having
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 PT or its analog.

XX Claim 11; Page 9; 43pp; English.

XX The invention relates to a synthetic immunogen for inducing specific

CC antibodies against gonadotrophin releasing hormone (GnRH also known as
 CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX

SQ Sequence 37 AA;

Query Match 100.0%; Score 29; DB 5; Length 37;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6

Db 11 SSGPSL 16

RESULT 7

AAU11430

ID AAU11430 standard; peptide; 46 AA.

XX AC AAU11430;

XX DT 12-MAR-2002 (first entry)

XX DE Synthetic immunogen peptide 11.

XX Gonadotrophin releasing hormone; GnRH; synthetic immunogen;

KW luteinising hormone releasing hormone; LHRH; contraceptive;

KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;

KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;

KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX Clostridium tetani.

OS Mammalia.

OS Synthetic.

OS Chimeric.

XX Key Location/Qualifiers

FT Peptide 1..10

FT /note= "Gonadotrophin releasing hormone epitope (1..10

FT aa)"

FT Misc-difference 1

FT /label= OTHER

FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"

FT Peptide 11..16

FT /note= "Spacer peptide"

FT Peptide 17..31

FT /note= "Tetanus toxoid (830-844 aa)"

FT Peptide 32..37

FT /note= "Spacer peptide"

FT Peptide 38..46

FT /note= "Gonadotrophin releasing hormone epitope (2-10

FT aa)"

FT Modified-site 46

FT /note= "Amidated glycine or glycylamide"

XX WO200185763-A2.

XX 15-NOV-2001.

XX 04-MAY-2001; 2001WO-US014363.

XX 05-MAY-2000; 2000US-0202328P.

XX (APHT-) APHTON CORP.

```

PI Grimes S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX
XX Novel synthetic immunogen for inducing immune response against
XX gonadotropin releasing hormone, comprises fusion peptide having
XX promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
XX or its analog.
XX
XX Claim 11; Page 12; 43pp; English.
XX
XX The invention relates to a synthetic immunogen for inducing specific
XX antibodies against gonadotropin releasing hormone (GnRH) also known as
XX luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
XX which comprises a promiscuous helper T-cell peptide epitope and
XX immunomimic peptide epitope or its analogue. The synthetic immunogen is
XX as such useful as a contraceptive and in the treatment of diseases
XX such as cancer (of the breast, uterus and other gynaecological cancer),
XX endometriosis, uterine fibroids, benign prostatic hypertrophy and
XX prostate cancer. The immunogen is effective in eliciting high and
XX specific anti-GnRH antibody titres. The present sequence is a synthetic
XX immunogen of the invention
XX
XX Sequence 46 AA;
XX
XX Query Match 100.0%; Score 29; DB 5; Length 46;
XX Best Local Similarity 100.0%; Pred. No. 1.3e+02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 SSGPSL 6
XX Db 11 SSGPSL 16
XX
XX RESULT 8
XX AAU11428
XX ID AAU11428 standard; peptide; 47 AA.
XX AC AAU11428;
XX
XX DT 12-MAR-2002 (first entry)
XX DE Synthetic immunogen peptide 9.
XX
XX KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;
XX luteinising hormone releasing hormone; LHRH; contraceptive;
XX promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
XX breast cancer; uterine cancer; gynaecological cancer; endometriosis;
XX uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
XX OS Plasmodium falciparum.
XX OS Mammalia.
XX OS Synthetic.
XX OS Chimeric.
XX
XX FH Key Location/Qualifiers
XX FT Peptide 1..10 /note= "Gonadotropin releasing hormone epitope (1..10
XX aa)"
XX FT Misc-difference 1 /label= OTHER
XX FT Peptide 11..16 /note= "Pyro-glutamic acid or 5-oxo proline"
XX FT Peptide 17..34 /note= "Spacer peptide"
XX FT Peptide 35..38 /note= "Malaria CSP protein (288-302 aa)"
XX FT Peptide 39..47 /note= "Spacer peptide"
XX FT Peptide /note= "Gonadotropin releasing hormone epitope (2-10
XX aa)"
XX FT Modified-site 47

```

```

FT /note= "Amidated glycine or glycylamide"
XX
XX WO200185763-A2.
XX
XX PD 15-NOV-2001.
XX
XX PF 04-MAY-2001; 2001WO-US014363.
XX
XX PR 05-MAY-2000; 2000US-0202328P.
XX
XX PA (APHT-) APHTON CORP.
XX
XX PI Grimes S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX
XX Novel synthetic immunogen for inducing immune response against
XX gonadotropin releasing hormone, comprises fusion peptide having
XX promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
XX or its analog.
XX
XX Claim 11; Page 11; 43pp; English.
XX
XX The invention relates to a synthetic immunogen for inducing specific
XX antibodies against gonadotropin releasing hormone (GnRH) also known as
XX luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
XX which comprises a promiscuous helper T-cell peptide epitope and
XX immunomimic peptide epitope or its analogue. The synthetic immunogen is
XX useful inducing an immune response against GnRH in an animal subject, and
XX as such is useful as a contraceptive and in the treatment of diseases
XX such as cancer (of the breast, uterus and other gynaecological cancer),
XX endometriosis, uterine fibroids, benign prostatic hypertrophy and
XX prostate cancer. The immunogen is effective in eliciting high and
XX specific anti-GnRH antibody titres. The present sequence is a synthetic
XX immunogen of the invention
XX
XX Sequence 47 AA;
XX
XX Query Match 100.0%; Score 29; DB 5; Length 47;
XX Best Local Similarity 100.0%; Pred. No. 1.3e+02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 SSGPSL 6
XX Db 11 SSGPSL 16
XX
XX RESULT 9
XX AAU11429
XX ID AAU11429 standard; peptide; 50 AA.
XX
XX AC AAU11429;
XX
XX DT 12-MAR-2002 (first entry)
XX DE Synthetic immunogen peptide 10.
XX
XX KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;
XX luteinising hormone releasing hormone; LHRH; contraceptive;
XX promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
XX breast cancer; uterine cancer; gynaecological cancer; endometriosis;
XX uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
XX OS Clostridium tetani.
XX OS Mammalia.
XX OS Synthetic.
XX OS Chimeric.
XX
XX FH Key Location/Qualifiers
XX FT Peptide 1..10 /note= "Gonadotropin releasing hormone epitope (1..10
XX aa)"
XX FT Misc-difference 1

```



```

FT FT /label= OTHER
FT FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT Peptide 11...16
FT Peptide /note= "Spacer peptide"
FT Peptide 17...36
FT Peptide /note= "Malaria CSP protein (378-398 aa)"
FT Peptide 37...42
FT Peptide /note= "Spacer peptide"
FT Peptide 43...51
FT Peptide /note= "Gonadotrophin releasing hormone epitope (2-10 aa)"
FT Modified-site 51
FT FT /note= "Amidated glycine or glycylamide"
XX WO200185763-A2.
XX 15-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US014363.
XX
XX 05-MAY-2000; 2000US-0202328P.
XX
XX (APHT-) APHTON CORP.
XX
XX Grimes S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX
XX Novel synthetic immunogen for inducing immune response against
XX gonadotropin releasing hormone, comprises fusion peptide having
XX promiscuous helper T-cell peptide epitope and immunomic peptide epitope
XX or its analog.
XX
XX Claim 11; Page 12-13; 43pp; English.
XX
XX The invention relates to a synthetic immunogen for inducing specific
XX antibodies against gonadotropin releasing hormone (GnRH) also known as
XX luteinising hormone releasing hormone (LHRH) comprising a fusion peptide
XX which comprises a promiscuous helper T-cell peptide epitope and
XX immunomic peptide epitope or its analogue. The synthetic immunogen is
XX useful inducing an immune response against GnRH in an animal subject, and
XX as such is useful as a contraceptive and in the treatment of diseases
XX such as cancer (of the breast, uterus and other gynaecological cancer),
XX endometriosis, uterine fibroids, benign prostatic hypertrophy and
XX prostate cancer. The immunogen is effective in eliciting high and
XX specific anti-GnRH antibody titres. The present sequence is a synthetic
XX immunogen of the invention
XX
XX Sequence 51 AA;
XX
Query Match 100.0%; Score 29; DB 5; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.4e+02; Mismatches 0; Gaps 0;
Matches 6; Conservative 0; Indels 0;
Oy 1 SSGPSL 6
Db 11 SSGPSL 16
RESULT 12
AA18164
ID AA18164 standard; protein; 53 AA.
XX
XX AA18164;
XX
XX 12-OCT-2001 (first entry)
XX
XX Peptide #4598 encoded by probe for measuring cervical gene expression.
XX
XX Probe; human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer.
XX
XX Homo sapiens.
XX

```

```

XX WO200157278-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000670.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human cervical epithelial cells.
XX
XX Claim 27; SEQ ID NO 22990; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs; see A110068-A128459). The present sequence is a peptide encoded
XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs
XX can be used to produce a single exon microarray, which can be used for
XX measuring human gene expression in a sample derived from human cervical
XX epithelial cells. By measuring gene expression, the probes are therefore
XX useful in grading and/or staging of diseases of the cervix, notably
XX cervical cancer. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 53 AA;
XX
Query Match 100.0%; Score 29; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 SSGPSL 6
Db 5 SSGPSL 10
RESULT 13
ABB37195
ID ABB37195 standard; peptide; 53 AA.
XX
XX ABB37195;
XX
XX 04-FEB-2002 (first entry)
XX
XX Peptide #4701 encoded by human foetal liver single exon probe.
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
XX Homo sapiens.
XX
XX WO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000669.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX

```


SQ Sequence 53 AA;

Query Match 100.0%; Score 29; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
| | | | |
Db 5 SSGPSL 10

Search completed: March. 10, 2004, 09:12:08
Job time : 11.035 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 5.92996 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-6
Perfect score: 29
Sequence: 1 SSGPSL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriaph.*
17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 29 | 100.0 | 113 | 16 | Q83F79 coxiella bu |
| 2 | 29 | 100.0 | 174 | 5 | Q8M201 drosophila |
| 3 | 29 | 100.0 | 200 | 3 | Q9UVT8 magnaporthe |
| 4 | 29 | 100.0 | 213 | 4 | Q9UG11 homo sapien |
| 5 | 29 | 100.0 | 213 | 11 | Q8VEL0 |
| 6 | 29 | 100.0 | 214 | 11 | Q9B8Y9 |
| 7 | 29 | 100.0 | 215 | 13 | Q7SYQ6 xenopus lae |
| 8 | 29 | 100.0 | 226 | 10 | Q9FUX4 |
| 9 | 29 | 100.0 | 257 | 16 | Q8VXE5 |
| 10 | 29 | 100.0 | 277 | 4 | Q8IXV0 |
| 11 | 29 | 100.0 | 284 | 2 | Q9L913 |
| 12 | 29 | 100.0 | 348 | 16 | Q8XUL9 |
| 13 | 29 | 100.0 | 358 | 16 | Q7U3U5 |
| 14 | 29 | 100.0 | 392 | 10 | Q7XK21 |
| 15 | 29 | 100.0 | 395 | 4 | Q8N4F2 |
| 16 | 29 | 100.0 | 395 | 12 | Q86606 simian viru |

| | | | | | |
|----|----|-------|------|----|--------|
| 17 | 29 | 100.0 | 405 | 16 | Q97H14 |
| 18 | 29 | 100.0 | 430 | 5 | Q7YVE4 |
| 19 | 29 | 100.0 | 437 | 16 | Q8XKM3 |
| 20 | 29 | 100.0 | 486 | 13 | Q7TOR3 |
| 21 | 29 | 100.0 | 492 | 16 | Q82M06 |
| 22 | 29 | 100.0 | 507 | 16 | Q8FY77 |
| 23 | 29 | 100.0 | 520 | 13 | Q7ZUA3 |
| 24 | 29 | 100.0 | 532 | 16 | P72605 |
| 25 | 29 | 100.0 | 593 | 16 | Q8A6C8 |
| 26 | 29 | 100.0 | 615 | 11 | Q8BP38 |
| 27 | 29 | 100.0 | 1009 | 16 | Q8PM63 |
| 28 | 29 | 100.0 | 1074 | 10 | Q8SYZ6 |
| 29 | 29 | 100.0 | 1197 | 16 | Q8NM57 |
| 30 | 29 | 100.0 | 2078 | 16 | Q98K31 |
| 31 | 29 | 100.0 | 3262 | 11 | Q9EQJ5 |
| 32 | 27 | 93.1 | 103 | 16 | Q89BU2 |
| 33 | 27 | 93.1 | 121 | 10 | Q9SV80 |
| 34 | 27 | 93.1 | 123 | 11 | Q9CWM5 |
| 35 | 27 | 93.1 | 183 | 5 | Q8SW89 |
| 36 | 27 | 93.1 | 189 | 11 | Q99PH9 |
| 37 | 27 | 93.1 | 189 | 11 | Q99PI3 |
| 38 | 27 | 93.1 | 192 | 10 | Q9SV78 |
| 39 | 27 | 93.1 | 207 | 10 | Q9FUS7 |
| 40 | 27 | 93.1 | 218 | 10 | Q9SXZ1 |
| 41 | 27 | 93.1 | 222 | 5 | O01786 |
| 42 | 27 | 93.1 | 227 | 10 | Q9FUS8 |
| 43 | 27 | 93.1 | 227 | 10 | Q94110 |
| 44 | 27 | 93.1 | 227 | 10 | Q9FUS9 |
| 45 | 27 | 93.1 | 227 | 11 | Q8VCG7 |

ALIGNMENTS

RESULT 1

Q83F79 PRELIMINARY; PRT; 113 AA.
AC Q83F79;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN CBU0069.
OS Coccidia burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coccidiaceae; Coccidia.
OX NCBI_taxid=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile phase I / RSA 493;
RX MEDLINE=22608657; PubMed=12704232;
RA Sehadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA Nelson M.C., Ward N.L., Tettelin H., Davidson T.M., Beanan M.J.,
RA DeBoy R.T., Daugherty S.C., Brinkac L.W., Madupu R., Dodson R.J.,
RA Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;
RT "Complete genome sequence of the Q-fever pathogen, Coccidia burnetii.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).
RL EMBL; AE016960; AAC09636.1; -.
DR TIGR; CBU0069; -.
DR InterPro; IPR002110; ANK.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 113 AA; 12752 MW; 5C82A677DEE7BD95 CRC64;

Query Match 100.0%; Score 29; DB 16; Length 113;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSGPSL 6
DB 35 SSGPSL 40

```

RESULT 2
Q8MZ01
ID Q8MZ01 PRELIMINARY; PRT; 174 AA.
AC Q8MZ01
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE RE33542p.
GN BCDNA:RE33542.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY113451; AAM29456.1; -
DR FLYBase; FBgn0063034; BCDNA:RE33542.
SQ SEQUENCE 174 AA; 19281 MW; 05C5CE1AB43DB8DA CRC64;

Query Match 100.0%; Score 29; DB 5; Length 174;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSSL 6
DB 122 SSGPSSL 127

RESULT 3
Q9UVT8
ID Q9UVT8 PRELIMINARY; PRT; 200 AA.
AC Q9UVT8
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE 6,7-dimethyl-8-ribityllumazine synthase.
OS Magnaporthe grisea (Rice blast fungus); (Pyricularia grisea).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
OX NCBI_TaxID=148305;
RN [1]
RP SEQUENCE FROM N.A.
RA Persson K., Schneider G., Jordan D.B., Viitanen P.V., Sandalova T.;
RA "Comparison of the crystal structures of the pentameric fungal and the
RT icosahedral plant lumazine synthases.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF148449; RAD55372.1; -
DR PDB; 1C41; 07-AUG-00
DR GO; GO:0003949; Criboflavin synthase complex; IEA.
DR GO; GO:0004746; Friboflavin synthase activity; IEA.
DR GO; GO:0009231; P-vitamin B2 biosynthesis; IEA.
DR InterPro; IPR002180; DMRL synthase.
DR Pfam; PF00885; DMRL synthase; 1.
DR ProDom; PD003664; DMRL synthase; 1.
DR TIGRFAMs; TIGR00114; ribH; 1.
SQ SEQUENCE 200 AA; 21072 MW; E079F049D32271B5 CRC64;

Query Match 100.0%; Score 29; DB 3; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSSL 6
DB 122 SSGPSSL 127

RESULT 4
Q9UUG1
ID Q9UUG1 PRELIMINARY; PRT; 213 AA.
AC Q9UUG1
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein dj473b4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rhodes S., Huckle E.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137163; CAB89662.1; -
DR EMBL; BC005700; AAH05700.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR005335; MSP domain.
DR Pfam; PF00635; MSP domain; 1.
DR PROSITE; PS50202; MSP; 1.
KW Hypothetical protein.
SQ SEQUENCE 213 AA; 24086 MW; 2E0944F4805B547F CRC64;

Query Match 100.0%; Score 29; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSSL 6
DB 155 SSGPSSL 160

RESULT 5
Q8VELO
ID Q8VELO PRELIMINARY; PRT; 213 AA.
AC Q8VELO
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE RIKEN cDNA 1810018L05 gene.
GN 1810018L05RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC018329; AAH18329.1; -
DR MGD; MGI:1917630; 1810018L05RIK.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR005335; MSP domain.
DR Pfam; PF00635; MSP domain; 1.
DR PROSITE; PS50202; MSP; 1.
SQ SEQUENCE 213 AA; 24074 MW; FB72756A0528CB34 CRC64;

Query Match 100.0%; Score 29; DB 11; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSSL 6
DB 155 SSGPSSL 160

```